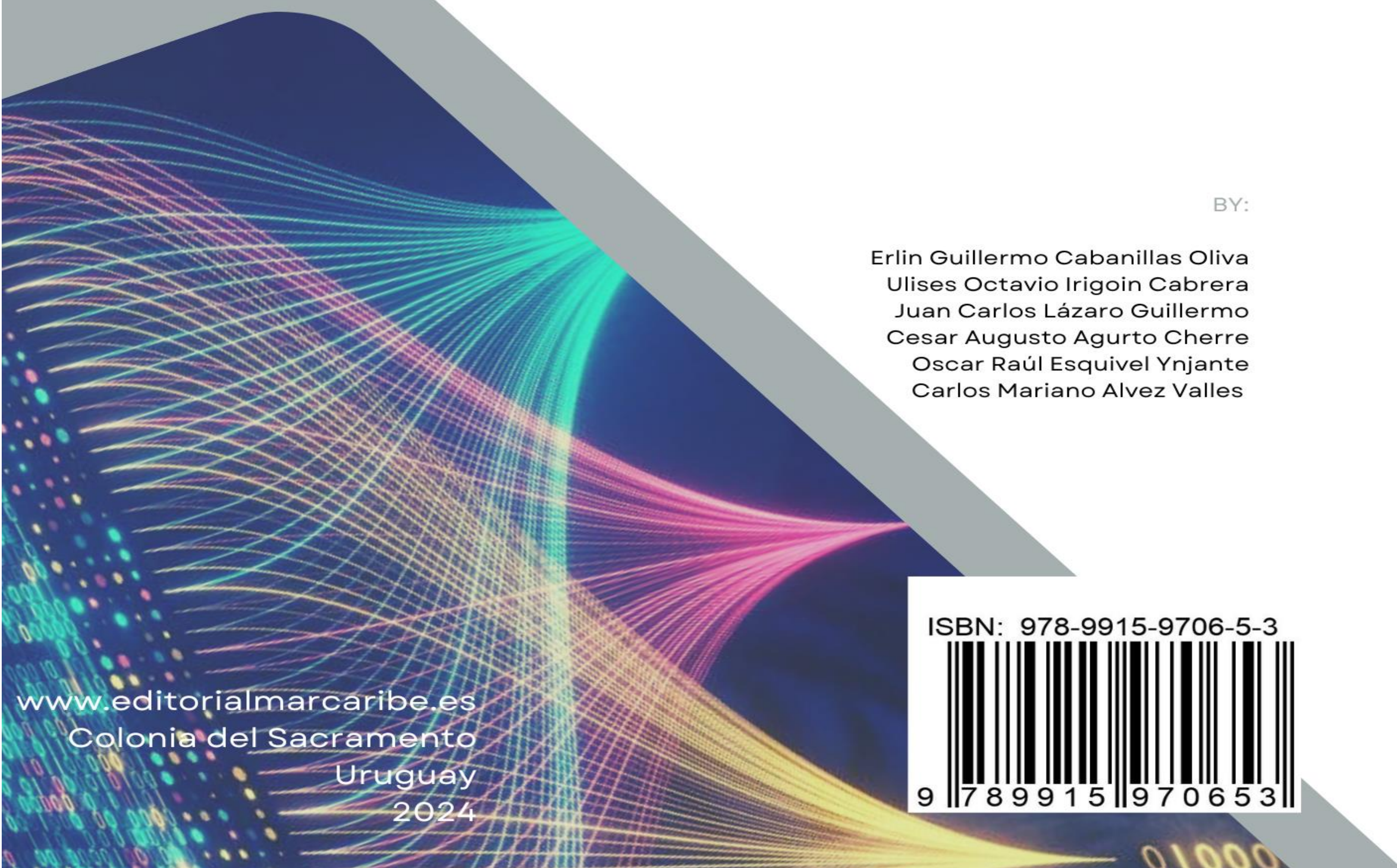
EMC
EST. 2021
EDITORIAL MAR CARIBE

DATA SCIENCE AND ARTIFICIAL INTELLIGENCE:
Finance, policy and governance

RESEARCH BOOK

BY:

Erlin Guillermo Cabanillas Oliva
Ulises Octavio Irigoin Cabrera
Juan Carlos Lázaro Guillermo
Cesar Augusto Agurto Cherre
Oscar Raúl Esquivel Ynjante
Carlos Mariano Alvez Valles



ISBN: 978-9915-9706-5-3

www.editorialmarcaribe.es
Colonia del Sacramento
Uruguay
2024

Data science and artificial intelligence: Finance, policy and governance

Erlin Guillermo Cabanillas Oliva, Ulises Octavio Irigoin Cabrera, Juan Carlos Lázaro Guillermo, Cesar Augusto Agurto Cherre, Oscar Raúl Esquivel Ynjante, Carlos Mariano Alvez Valles

Second edition: September, 2024

Edited by:

Editorial Mar Caribe

www.editorialmarcaribe.es

Av. General Flores 547, Colonia, Colonia-Uruguay.

Cover design: Yelitza Sanchez Caceres

Translation of the original Spanish edition into English: Ysaelen Josefina Odor Rossel

E-book available at https://editorialmarcaribe.es/data-science-and-artificial-intelligence-finance-policy-and-governance/

Format: electronic

ISBN: 978-9915-9706-5-3

ARK: ark:/10951/isbn.9789915970653

Editorial Mar Caribe

Data science and artificial intelligence: Finance, policy and governance

Colonia, Uruguay

2024

About the authors and the publication

Erlin Guillermo Cabanillas Oliva

egco1egco1@gmail.com

https://orcid.org/0000-0001-9815-6828

Universidad Nacional de la Amazonía Peruana, Peru

Ulises Octavio Irigoin Cabrera

uirigoin@ucp.edu.pe

https://orcid.org/0009-0007-6168-7415

Universidad Científica del Perú, Peru

Juan Carlos Lázaro Guillermo

jlazarog@unia.edu.pe

https://orcid.org/0000-0002-4785-9344

Universidad Nacional Intercultural de la Amazonía, Peru

Cesar Augusto Agurto Cherre

cesar_agurto@unu.edu.pe

https://orcid.org/0000-0001-6494-3567

Universidad Nacional de Ucayali, Peru

Oscar Raúl Esquivel Ynjante

oesquiveli@unia.edu.pe

https://orcid.org/0000-0002-5097-831X

Universidad Nacional Intercultural de la Amazonía, Peru

Carlos Mariano Alvez Valles

calvezv@unmsm.edu.pe

https://orcid.org/0000-0003-2341-6191

Universidad Nacional Mayor de San Marcos, Peru

Research result book:
Original and unpublished publication, the content of which is the result of a research process carried out before its publication, has been double-blind peer-reviewed by external peers, the book has been selected for its scientific quality and because it contributes significantly to the area of knowledge and illustrates fully developed and completed research. In addition, the publication has undergone an editorial process that guarantees its bibliographic standardization and usability.

Index

Introduction

Incorporating artificial intelligence (AI) and big data into sentiment analysis to detect patterns, trends and trading signals is a growing trend that has been around for some time. For years, traders have carefully analyzed news and management reports, trying to understand how non-financial information affects stock prices (Assad et al., 2020).

However, the use of advanced technologies such as text mining, social network analysis, and natural language processing (NLP) algorithms has taken this method to a new level. These innovative tools allow marketers to make informed decisions by automating data collection and analysis, as well as identifying consistent patterns or behaviors on a scale that humans cannot handle (Schrepel, 2020).

Therefore, AI-powered trading differs from systematic trading in its use of reinforcement learning and its ability to fine-tune the AI model to changing market conditions. In contrast, traditional methodological strategies often take longer to fine-tune parameters due to extensive human involvement. Traditional back testing strategies based on historical data may not produce optimal real-time performance when pre-established trends are no longer valid. On the other hand, the implementation of machine learning models allows analytics to focus on predicting and analyzing trends in real-time.

These tests predict and adapt to trends in real-time, thereby minimizing the risk of overfitting or curve-fitting seen in back testing based solely on historical data and trends. The application of artificial intelligence in trading has gone through many stages of development and has become increasingly complex, integrating at each stage with traditional algorithmic trading. Initially, algorithms were simple, with predefined buy or sell orders and basic parameters.

Later, more advanced algorithms were introduced that allowed for flexible pricing. The next generation of algorithms focuses on minimizing market impact by splitting large orders, called "execution algorithms," to achieve optimal pricing. Today, advanced strategies use deep neural networks to optimize order placement and execution, with the goal of minimizing market impact (Tan, 1997). Inspired by the human brain, deep neural networks use algorithms that can recognize patterns and require less human intervention to operate and learn. Using these techniques, market makers can improve inventory management and reduce balance sheet costs.

As artificial intelligence continues to develop, algorithms are moving towards automation, relying more on computer programming and learning from input data, thereby reducing the need for human intervention. In practical applications, more advanced forms of AI are currently used primarily to detect signs of trouble in flow-based trading that may not have much news value. These incidents are less visible, pose greater challenges to identify, and extracting value from them is a more difficult task.

Rather than simply improving execution speed, AI is actually used to filter out data noise and turn that information into actionable decisions. On the other hand, less complex algorithms are mostly used for information-rich events, such as financial news, which are easy for all participants to understand and require rapid implementation.

Therefore, at the current stage of development, ML-based models serve a different purpose than HFT strategies, which focus on acting quickly and gaining an edge in trading. Instead, ML models are mostly used offline for tasks such as fine-tuning algorithm parameters and optimizing decision logic rather than performing actual trades.

However, as AI technology advances and its applications increase, it has the potential to improve traditional algorithmic trading in the future (OECD, 2023).

This is possible when AI technologies are integrated into the trade execution phase, providing advanced automated trade execution features and covering every stage from signal collection to strategy and trade execution.

ML-based execution algorithms will enable automatic and dynamic adjustment of decision logic during trading. In such cases, current requirements for algorithmic trading, such as safeguards in pre-trade risk management systems and automated control mechanisms to stop algorithms when they exceed risk limits, should be expanded to include AI-guided algorithmic trading.

Chapter 1

Artificial Neural Networks: Financial risks in credit institutions

Since the groundbreaking research conducted by Beaver in the late 1960s, there has been a great deal of interest in using financial ratios as a means of predicting financial failure. This surge in interest can be attributed to the influential work of Altman (1968), where he combined five financial ratios into a single predictor known as the Z factor, specifically designed to assess the likelihood of business failure (Tan, 1997). A notable advantage of Altman's methodology is its ability to establish a standard benchmark for comparing companies within the same industry, while providing a consolidated measure of financial strength derived from a company's financial accounts. However, despite its appeal, this methodology is not without limitations, as ratios can vary significantly across different industry sectors and accounting methods used.

Limitations become more apparent when using financial indicators to forecast the financial challenges faced by financial institutions. The inherent high leverage of these institutions makes it difficult to apply models that were originally developed for the corporate sector. However, there has been increasing acceptance of using these models in the financial sector by considering financial institutions as a distinct category of firms. In Australia, there have been instances where researchers have conducted unprecedented analyses of financial distress among non-bank financial institutions. These studies employ a Probit model to address the limited nature of the dependent variables observed in the financial distress data.

In this section of the book, the main focus is on the effectiveness of ANNs as an early indicator of financial distress within credit unions. To provide an unbiased

assessment, the ANN-based model developed in this study is compared to the Probit model created by Hall and Byron, using the same data set. The findings suggest that the ANN method slightly outperforms the Probit model when examining the same data set. In addition, modifications to the ANN model design are explored to improve its performance as an early warning predictor.

Benefits of AI in Finance

The adoption of artificial intelligence (AI) in the financial industry is being driven by the significant and ever-increasing availability of data, as well as the advantage that AI and machine learning (ML) can bring to financial services firms (OECD, 2021). With the explosion of data and advances in computing power, particularly through cloud computing, machine learning models can effectively analyze this vast amount of data and uncover hidden patterns and relationships that are beyond human capabilities.

As a result, financial sector companies are increasingly using AI/ML and big data to gain a competitive advantage. This includes improving operational efficiency by reducing costs and improving the quality of financial services products to meet customer demands. This trend is expected to further amplify the competitive advantage of financial companies in the future (OECD, 2021).

Portfolio and asset management

ML models have the ability to continuously monitor and analyze thousands of risk factors on a daily basis. Additionally, they can simulate and evaluate portfolio performance under thousands of economic and market scenarios. This level of advanced analysis and risk assessment can improve risk management practices for asset managers and other large institutional investors. One specific application of AI, known as natural

language generation (NLG), can prove especially valuable to financial advisors. NLG enables advisors to analyze and present complex data in a more understandable and relatable way for their clients. By "humanizing" and simplifying data analysis and reporting, NLG can help advisors effectively communicate investment strategies and insights to their clients.

Thus, the use of AI and ML in asset management offers a multitude of benefits. From improving operational efficiency to enhancing risk management practices and delivering a superior customer experience, these technologies have the potential to revolutionize the industry. As the field of AI continues to advance, asset managers and financial institutions are expected to increasingly adopt these technologies to stay ahead in an ever-evolving market. In terms of operational benefits, the implementation of AI technologies can result in significant cost reductions for investment managers.

By automating tasks that were previously performed manually, such as reconciliation processes, AI can streamline operations and reduce administrative expenses. Additionally, the increased efficiency and speed offered by AI can potentially lead to greater cost savings for asset managers. The integration of artificial intelligence (AI) and machine learning (ML) into asset management has the potential to improve the efficiency and accuracy of various operational workflows. This technological advancement not only promises to improve overall performance but also strengthen risk management practices and enhance the overall customer experience.

By using large amounts of data, machine learning models can offer asset managers valuable recommendations that can impact their decision-making process regarding portfolio allocation and stock selection. With the advent of big data, traditional data sets have become widely accessible to all investors, prompting asset managers to leverage this resource to gain valuable insights into their investment strategies.

Across the investment community, information has always played a crucial role, with data serving as the foundation for various investment approaches such as fundamental analysis and systematic trading. While structured data has long been the focal point of these “traditional” strategies, the abundance of raw or unstructured/semi-structured data now presents an opportunity for investors to use AI to gain a new informational advantage. By employing AI, asset managers can efficiently process large amounts of data from multiple sources and quickly extract valuable insights to inform their strategies.

The use of artificial intelligence and machine learning, along with big data analytics, tends to be more common among large asset managers and institutional investors due to their financial capacity and available resources to invest in AI technologies. As a result, smaller players may face difficulties in adopting these techniques, as they lack the necessary investment in technology and skilled professionals to handle large amounts of unstructured big data and develop machine learning models.

Even if the implementation of AI and proprietary models provides a competitive advantage, it may further limit the participation of smaller players who cannot incorporate in-house AI techniques or access big data sources (França et al., 2021). This could therefore reinforce the current trend of concentration among a few large players in the hedge fund sector, as these larger groups outperform their more agile competitors.

Limited participation by smaller entities in the sector will continue until the tools they need are widely available or offered by third-party vendors. In addition, third-party data sets may not meet the same industry standards, so users of these tools will need to build trust in the accuracy and reliability of the information they rely on. This level of confidence in the validity of big data is necessary for smaller players to feel comfortable enough to adopt and use these tools.

The use of identical AI models across multiple asset managers has the potential to lead to herding behavior and create one-way markets. This could present certain dangers to the overall liquidity and stability of the system, particularly during periods of economic stress. The emergence of significant market volatility may be intensified by simultaneous large-scale buying or selling activities, thereby introducing new vulnerabilities into the system.

There is a possibility that incorporating AI/ML and big data into investment strategies has the potential to reverse the prevailing trend of passive investing. If these innovative technologies demonstrate a consistent ability to generate alpha, indicating a cause-effect relationship between the use of AI and outperformance, it presents an opportunity for the active investment community to reinvigorate its approach and provide additional alpha opportunities to its clients.

AI-powered hedge funds and ETFs

Hedge funds have been leading the way in adopting and utilizing innovative financial technology, such as big data analytics, artificial intelligence (AI) and machine learning (ML), in their trading strategies and back-office operations. In more recent times, a new generation of hedge funds, commonly referred to as “AI pure play” funds, has emerged that rely exclusively on AI and ML technologies to drive their investment decisions and portfolio management (e.g., Aidiyia Holdings, Cerebellum Capital, Taaffeite Capital Management and Numerai).

So far, there has been a notable absence of any academic or impartial assessment of the effectiveness of artificial intelligence (AI)-powered funds, conducted by an entity outside the financial industry. Such an assessment would aim to compare the numerous funds that claim reliance on AI technology (Westerhuis et al., 2008). As fund managers

employ varying levels of AI integration in their operations and strategies, they naturally hold back their methodologies to maintain a competitive edge. Consequently, it becomes difficult to compare the performance of various self-proclaimed AI-powered products, as the degree of AI utilization and the maturity of its implementation differ significantly across these funds (Motta, 2023).

The private sector offers AI-powered hedge fund indices that clearly outperform conventional hedge fund indices provided by the same source. It is important to note that third-party indices are often influenced by biases such as survivorship bias and self-selection of funds included in the index, as well as backfilling. It is therefore advisable to approach these indices with caution.

Furthermore, there is growing evidence suggesting that machine learning (ML) models outperform traditional forecasts when it comes to macroeconomic indicators such as inflation and GDP. This improvement in performance is particularly evident in times of economic stress when accurate forecasts are crucial. Thus, AI-based techniques have proven superior in identifying previously unknown correlations in the occurrence of financial crises. ML models have significantly outperformed logistic regression models in predicting and forecasting financial crises in out-of-sample tests.

Algorithm trading

Artificial intelligence has the potential to revolutionize the trading industry by offering trading strategy suggestions and powering automated trading systems. These AI-based systems are capable of making predictions, determining the best course of action, and executing trades without the need for human intervention. They use advanced AI techniques such as evolutionary computing, deep learning, and probabilistic logic to identify and execute trades in the market.

Similarly, AI techniques such as algorithmic wheels can systematically strategize upcoming trades by applying a logical "if/then" thought process. This level of AI integration into trading enables predictive capabilities that far exceed those of traditional algorithms in the financial and trading sectors, particularly considering the current interconnectedness across asset classes and geographies.

Thus, AI-powered trading systems have the potential to assist traders in effectively managing both their risk and order flow. These innovative applications can monitor and analyze risk exposure, allowing them to automatically adjust or exit positions based on user preferences and requirements. The remarkable aspect of these AI systems is that they possess the ability to self-train and adapt to ever-changing market conditions, thereby minimizing the need for human intervention. Likewise, these systems can facilitate the seamless management of flows between brokers, ensuring smooth execution of predetermined trades. And also, they have the ability to regulate fees and allocate liquidity across various exchanges, considering factors such as regional market preferences, monetary considerations, and other essential parameters involved in managing an order.

In today's technologically advanced markets, particularly in the fields of equity and foreign exchange products, the implementation of AI solutions has great potential in terms of providing competitive pricing, efficient liquidity management, and optimized execution processes (Botta & Wiedemann, 2019). One of the crucial advantages of using AI algorithms in trading is their ability to improve liquidity management and facilitate the execution of large orders without causing substantial market disruptions. These algorithms possess the ability to dynamically adjust order size, duration, and order size, based on the prevailing market conditions, thereby ensuring optimal performance.

The integration of artificial intelligence (AI) and big data into sentiment analysis to detect patterns, trends, and trading signals is a growing trend that has been around for quite some time now. Traders have been examining news and statements from company management for years, trying to understand how non-financial information affects stock prices. However, the use of advanced technologies such as text mining, social media analysis, and natural language processing (NLP) algorithms has taken this practice to new heights. These innovative tools allow traders to make informed decisions by automating the process of data collection and analysis, as well as identifying consistent patterns or behaviors on a scale that would be impossible for a human to handle.

Consequently, AI-driven trading is distinguished from systematic trading due to its utilization of reinforcement learning and the ability to adjust the AI model according to changing market conditions. In contrast, traditional systematic strategies often require more time to fine-tune parameters due to extensive human involvement. Conventional back testing strategies, which are based on historical data, might not deliver optimal performance in real-time when previously identified trends no longer hold. On the other hand, the implementation of machine learning models allows the analysis to focus on predicting and analyzing trends in real-time. For example, predictive testing is employed instead of back testing. These tests predict and adapt to trends in real-time, thereby mitigating the risk of overfitting or curve-fitting observed in back testing based solely on historical data and trends.

The application of AI in trading has gone through several phases of development and increasing complexity, integrating with traditional algorithmic trading at each stage. Initially, algorithms were simple, with predefined buy or sell orders and basic parameters. Later, more advanced algorithms were introduced that allowed for dynamic pricing (Brown & MacKay, 2021). The next generation of algorithms focused on

minimizing the impact on the market by breaking up large orders, known as "execution algorithms," which aimed to obtain optimal prices.

Currently, innovative strategies use deep neural networks to optimize order placement and execution, with the goal of minimizing market impact. Deep neural networks, inspired by the human brain, employ algorithms that are capable of recognizing patterns and require less human intervention to operate and learn. By using these techniques, market makers can improve their inventory management and reduce balance sheet costs. As AI continues to advance, algorithms are moving toward automation, relying more on computer programming and learning from input data, thereby reducing the need for human intervention (Metaxa et al., 2021).

In the realm of practical application, the most advanced forms of AI are currently predominantly used to detect incident signals in flow-based trading that may not have significant news value. These incidents are characterized by being less overt, posing greater challenges in identification, and extracting value from them is a more arduous task. Rather than merely improving execution speed, AI is actually employed to filter out data noise and transform this information into actionable decisions. On the other hand, less sophisticated algorithms are employed in information-laden events, such as financial news, which are more easily understandable to all participants and require fast execution.

Therefore, at the current stage of their development, ML-based models serve a different purpose compared to HFT strategies, which focus on quick action and gaining an edge in trading. Instead, ML models are mostly used offline for tasks such as refining algorithm parameters and improving decision-making logic rather than for actual trade execution. While, as AI technology advances and finds more applications, it has the potential to improve traditional algorithmic trading in the future. This could happen when AI techniques are incorporated into the trade execution phase, providing advanced

capabilities for automated trade execution and covering every step from signal capture to trade strategy and execution.

ML-based execution algorithms would enable autonomous and dynamic adjustment of decision logic during trading. In such cases, existing requirements for algorithmic trading, such as safeguards in pre-trade risk management systems and automated control mechanisms to stop algorithms when they exceed risk limits, would need to be extended to include AI-powered algorithmic trading (OECD, 2023).

The unintended consequences and potential risks of AI

The widespread adoption of identical or similar models by numerous operators in various markets may have unintended consequences for competition and could exacerbate tensions within those markets (Descamps et al., 2021). If these models were to become widely used by traders, they would naturally decrease arbitrage opportunities, resulting in lower profit margins. However, this would benefit consumers, as it would reduce the difference between buying and selling prices.

On the other hand, it could also lead to market convergence, where traders follow the same strategies, creating a herd mentality and causing markets to move in only one direction. This could potentially affect market stability and liquidity, especially during times of high stress. Like any algorithm, extensive use of similar AI algorithms carries the risk of self-reinforcing feedback loops, which can trigger major price fluctuations.

Furthermore, the use of AI in malicious activities has the potential to lead to offensive autonomous attacks. These attacks can be carried out without human intervention, making them even more dangerous. Not only can vulnerable systems in trading be targeted, but financial markets as a whole, including the various participants

in them, are also at risk. This highlights the wide scope of the potential impact of AI-based cyberattacks.

The convergence of AI technologies not only enhances cybercriminals' capabilities to exploit interconnected systems but also enables the execution of autonomous attacks. These attacks can have serious consequences for both commerce and financial markets, requiring increased cybersecurity measures to mitigate the risks. The convergence of AI technologies presents not only opportunities but also risks, particularly in the area of cyberattacks. As AI systems become more interconnected and unified in their actions, cybercriminals can exploit this unity to their advantage. They find it easier to manipulate and influence agents that share similar behaviors than those with distinct and differentiated behaviors. This convergence poses a significant threat to cybersecurity.

The use of proprietary models that cannot be replicated plays a crucial role in allowing operators to maintain any form of competitive advantage. Furthermore, these proprietary models can contribute to a deliberate lack of transparency, thus exacerbating the challenge of understanding and explaining machine learning models. The reluctance shown by users of machine learning techniques to disclose the effectiveness of their models is due to the fear of compromising their competitive advantage, which in turn raises concerns regarding the oversight of machine learning algorithms and models (OECD, 2021).

The use of algorithms in trading can also facilitate and increase the likelihood of collusive outcomes in digital markets (Botta & Wiedemann, 2020). Furthermore, there are concerns that AI-based systems may worsen illegal practices aimed at manipulating markets, such as "spoofing," by creating difficulties for regulators to detect such activities when machines collude. The lack of explainability of the machine learning models used to support trading can pose difficulties in adjusting strategies during periods of poor

trading outcomes. Trading algorithms no longer follow linear, model-based processes (where input A leads to the execution of trading strategy B) that can be easily traced and interpreted, making it less clear which parameters influenced the outcomes.

When considering the potential negative consequences on the market, it can be argued that the use of AI technologies in trading and high-frequency trading (HFT) could potentially intensify market volatility by executing large simultaneous buys or sells. This introduces new vulnerabilities within the market. In particular, certain algo-HFT strategies have been implicated in the emergence of extreme market volatility, decreased liquidity, and exacerbated flash crashes, which have become more frequent in recent years. Since HFTs play an important role in providing liquidity to the market and improving its efficiency under normal conditions, any disruption in the operation of their models in times of crisis may lead to a withdrawal of liquidity from the market, potentially affecting its resilience.

In the investment arena, the widespread use of pre-existing AI models by multiple market participants has the potential to have a major impact on market liquidity and stability. This impact arises from the tendency of these models to encourage herding and one-way markets. Such behavior not only magnifies the risks associated with volatility, procyclicality, and unforeseen market swings, but also affects the scale and direction of the market. Furthermore, herding behavior can lead to illiquid markets if there are no "buffers" or market makers present to transact from the opposite side.

The introduction of AI into trading has the potential to create unforeseen connections between financial markets and institutions, leading to increased correlation and dependence of previously unrelated variables. The use of algorithms that generate profits or returns without any correlation may actually result in the correlation of unrelated variables if their use becomes widespread enough. Furthermore, the use of AI

can magnify the impact of network effects, leading to unexpected changes in the size and direction of market movements.

To address the risks associated with implementing AI in trading, it may be necessary to put safeguards in place for AI-driven algorithmic trading. These safeguards, built into pre-trade risk management systems, are designed to prevent and stop potential misuse of these systems. It should be noted that AI is also being used to improve pre-trade risk systems, encompassing mandatory testing of each algorithm version, which applies equally to those based on AI. As a final defense for market professionals, automated control mechanisms are in place to immediately shut down the model when it exceeds the limits of the risk system. These mechanisms involve "pulling the plug" and replacing any technology with human intervention. However, such measures may be considered suboptimal from a policy perspective, as they take systems offline precisely when they are most needed in times of stress and create operational vulnerabilities.

In addition to implementing safeguards on the exchanges where transactions take place, it may be imperative to employ various defensive measures. These measures could involve automatically cancelling orders whenever the AI system experiences an offline state, as well as employing techniques that provide resilience against sophisticated forms of manipulation facilitated by technology. There is also the possibility of modifying circuit breakers, which are currently triggered by significant drops in trading, to also recognize and trigger in response to a substantial volume of smaller trades executed by AI-powered systems, thereby achieving a similar outcome.

AI: "something wheels"

Something wheels refers to a broad concept that includes fully automated solutions designed to guide trader-driven flow. In this context, an AI-based algorithmic wheel

represents an automated routing process that integrates artificial intelligence techniques to assign a suitable broker algorithm to orders from a predetermined list of algorithmic solutions (Cheng & Nowag, 2023). AI-based algorithm wheels serve as models that determine the most advantageous strategy and broker to route the order, considering the prevailing market conditions as well as the specific objectives and requirements of the trading activity.

Investment firms typically use algorithmic trading wheels for two main purposes:

- First, they use these wheels to improve performance by achieving better quality of execution.
- Secondly, they leverage algorithm wheels to streamline their workflow by automating the handling of smaller orders and implementing standardized naming conventions for brokers’ algorithms.

Proponents of algorithm wheels argue that they effectively mitigate traders' bias when it comes to choosing brokers and the algorithms they implement in the market.

According to recent estimates, 20% of trade flows currently use algo wheels, a mechanism that is gaining popularity to systematically categorize and measure the effectiveness of algorithms used by high-performing brokers. Interestingly, those that employ algo wheels allocate a significant 38% of their trade flow to this tool (OECD, 2021). This suggests that if algorithmic wheels were to be widely adopted, it could lead to a significant increase in the overall volume of e-trading, which in turn could lead to various advantages for the e-brokerage competitive landscape.

Chapter 2

BigTech, financial services and blockchain

As tech giants continue to use their unrestricted access to vast amounts of customer data to power AI-driven systems to deliver financial services, there is a growing need to examine the data privacy implications of their deployment of AI. This has raised concerns regarding the potential exploitation of the collection, storage and utilization of personal data for commercial gain. The practices employed by BigTechs in this regard have the potential to negatively impact customers, particularly through discriminatory practices affecting credit availability and pricing.

BigTechs' access to customer data gives them a significant advantage over traditional financial services providers. This advantage is expected to be further strengthened as they incorporate artificial intelligence into their services, enabling the delivery of unique, personalized and more efficient offerings. However, BigTechs' dominance in certain areas of the market can lead to over-concentration and increased reliance on a small number of large players.

Depending on the size and scope of these companies, this could have systemic implications and raise concerns about potential risks to financial consumers. These consumers may not have access to the same range of product options, pricing, or advice that would be available through traditional financial services providers. And supervisors may face challenges in monitoring and regulating the financial activities of these big tech companies.

Another risk related to this issue is the potential for anti-competitive behavior and concentration of market power in the technological field of service delivery. This could

occur if only a few dominant players emerge in the markets for AI solutions and services using AI technologies. This trend is already being observed in certain regions of the world. The competitive landscape is also further compromised by the advantageous position held by large technology companies in terms of customer data (Butijn, 2023). These companies can exploit their data advantage to establish monopolistic positions, gaining an advantage in customer acquisition through effective price discrimination and creating significant barriers to entry for smaller companies.

Overall, the Digital Markets Act represents a significant step towards regulating and supervising the activities of dominant digital platforms in order to promote fair competition and protect the interests of business users. The Digital Markets Act includes a number of obligations that gatekeepers would be required to comply with. One of those obligations is to provide access to data generated by their activities to business users. Gatekeepers would also be required to offer data portability, allowing users to easily transfer their data to other platforms. To prevent unfair competition, gatekeepers would also be prohibited from using data obtained from business users to compete against them.

In late 2020, the European Union and the United Kingdom jointly published a set of regulatory proposals known as the Digital Markets Act. These proposals are designed to create a proactive framework for regulating dominant digital platforms, commonly known as "gatekeepers," such as large technology companies. The main goal of these proposals is to address the risks associated with these platforms and establish fair and open digital markets.

Another key aspect of the proposal is the introduction of measures to address the risks associated with gatekeepers' dual roles. This would involve implementing solutions to address issues such as self-referral, where gatekeepers prioritize their own services over those of third parties. The proposal also aims to ensure that services offered by

gatekeepers are not given preferential treatment or favored over those provided by third-party platforms.

AI and Blockchain-based financial products

In recent years, there has been a significant increase in the utilization of distributed ledger technologies (DLT), particularly blockchain, across various sectors, with a strong focus on the financial industry. This rise in blockchain applications can be attributed to the numerous benefits they offer, such as increased speed, efficiency, and transparency. These innovative technologies, driven by automation and disintermediation, have gained traction due to their potential to revolutionize different areas, including securities markets (such as issuance and post-trade activities), payments (such as digital currencies and central bank stablecoins), and asset or tokenization in general. As a result, the adoption of DLT in finance has the potential to reshape the functions and business models of financial operators, such as custodians.

The industry is advocating for the fusion of artificial intelligence (AI) and distributed ledger technology (DLT) in the realm of blockchain-based finance. This integration is believed to improve the overall effectiveness of these systems by leveraging automation to maximize the efficiency promised by blockchain-based solutions. However, it is currently unclear whether the scope of AI implementation in blockchain-based projects is substantial enough to substantiate claims of a true convergence between these two technologies.

In practice, rather than seeing convergence, what we see is the implementation of AI applications on specific blockchain systems for particular use cases, such as risk management. We also witness the integration of DLT solutions into certain AI mechanisms, particularly for data management purposes. Integration involves the use of

DLT to provide input to machine learning models, taking advantage of the blockchain's immutability and disintermediation features.

The integration undoubtedly enables the secure sharing of sensitive information while maintaining confidentiality and privacy. It is anticipated that the incorporation of DLT into AI mechanisms will enable users to monetize the data they possess, which is used by machine learning models and other AI-driven systems such as the Internet of Things (IoT) (Moujahid, 2016). The adoption of these AI use cases is motivated by the technology's potential to improve automation efficiency and disintermediation in DLT-based systems and networks.

Artificial intelligence has the potential to significantly improve the automation capabilities of smart contracts in the realm of DLT-based finance. This contribution can be seen in numerous use cases within DLT networks, including but not limited to compliance and risk management. For example, AI can play a crucial role in combating fraudulent activities by implementing automated restrictions on the network. Furthermore, AI can also improve the functioning of oracles, which are essential for data management and inference. However, it is important to note that these applications are still under development and refinement.

AI has the potential to significantly improve the security and functionality of blockchain networks, particularly in the realm of payment applications. While it cannot completely eliminate security vulnerabilities, AI can effectively mitigate them. By leveraging AI technology, users of blockchain networks can detect and address irregular activities that may be indicative of theft or fraudulent behavior, although these events typically require the compromise of public and private keys.

Similarly, AI applications can play a critical role in streamlining onboarding procedures, such as using biometrics for AI-based identification and strengthening anti-money laundering and countering the financing of terrorism (AML/CFT) controls for DLT-based financial services. Integrating AI into DLT-based systems also offers significant benefits in terms of compliance and risk management. For example, AI-powered tools can generate portfolio governance analysis results, which can serve as valuable resources for compliance purposes or internal risk assessments of transaction parties. However, it is important to recognize that removing financial intermediaries from financial transactions can undermine the effectiveness of existing regulatory approaches that focus primarily on regulated entities.

By incorporating AI-powered solutions into distributed ledger technology (DLT) systems at the protocol level, regulators can effectively achieve their regulatory objectives. This can be achieved through several means, such as facilitating seamless, real-time data exchange between entities and regulated authorities, as well as incorporating regulatory requirements directly into program code to ensure automatic compliance. The concept of regulators becoming nodes in decentralized networks has been a topic of discussion within the market, as it presents a potential solution to the challenges of overseeing platforms that operate without a central authority.

In relation to data in DLT-based systems, AI has the potential to enhance these processes. By shifting the responsibility for data curation from third-party nodes to independent, automated AI-powered systems, the quality of data fed into the chain can be significantly improved. This, in turn, leads to more robust recording and sharing of information, as AI systems are less prone to manipulation. One area where AI can make a particular difference is in the operation of third-party off-chain nodes, commonly known as "Oracles," which play a crucial role in feeding external data into the network.

The use of oracles in distributed ledger technology (DLT) networks exposes a potential risk of erroneous or inappropriate data being entered into the network, arising from the possibility of malicious or poorly performing third-party off-chain nodes. To address this issue, integrating artificial intelligence (AI) on-chain could improve disintermediation by making third-party information providers, such as oracles, unnecessary.

By incorporating AI, the system can verify the accuracy and integrity of the data provided by the oracles, thus preventing cyberattacks and manipulation of third-party data supply within the network. Implementing AI applications could potentially improve the trust of participants in the network, as they can verify the information provided by the oracles and identify any compromise within the system. However, it is important to note that AI does not inherently solve the problem of poor quality or inadequate input data, as this challenge is also present in AI-based mechanisms and applications.

AI increases the capabilities of smart contracts

The integration of AI techniques into blockchain-based systems has the potential to bring about significant changes, particularly in the realm of smart contracts. This integration may have practical implications on the governance and risk management of these contracts and may introduce several hypothetical, yet-to-be-tested effects on the functions and processes of distributed ledger technology (DLT)-based networks. Using AI in this context may pave the way for self-regulating DLT chains that operate autonomously.

Smart contracts have been around for a considerable time, predating the emergence of AI applications, and operate with simple software code. Currently, most widely used smart contracts do not incorporate AI methods. Consequently, numerous

proposed advantages of incorporating AI into DLT systems remain primarily theoretical, and it is advisable to approach claims made by industry regarding the integration of AI and DLT capabilities into commercialized products with a sense of careful consideration.

That said, the use of AI in various scenarios is highly advantageous when it comes to improving the capabilities of smart contracts, particularly in the areas of risk management and identifying flaws within the code. AI methodologies such as natural language processing (NLP) can effectively assess the execution patterns of smart contracts, thereby enabling the detection of fraudulent activities and improving the overall security of the system.

AI also has the ability to perform code testing in a way that surpasses the capabilities of human code reviewers in terms of speed and level of detail, as well as scenario analysis. Since code forms the fundamental basis of smart contract automation, the flawless nature of the coding process is crucial to ensuring the resilience and reliability of these contracts.

The potential to enhance the automation capabilities of smart contracts by integrating AI is a promising concept. By incorporating AI into smart contracts, the level of autonomy can be increased, allowing the underlying code to dynamically adapt to changing market and environmental conditions. A specific area of AI, known as natural language processing (NLP), has the potential to expand the analytical capabilities of smart contracts, particularly in relation to traditional contracts, legislation, and court rulings. By leveraging NLP, smart contracts can gain deeper insights into the intentions of the parties involved. While it is worth mentioning that these applications of AI for smart contracts remain purely theoretical and have yet to be tested in real-world scenarios.

There are still challenges that need to be addressed when it comes to operational risks, as well as compatibility and interoperability between traditional infrastructure and one based on DLT and AI technologies. The use of AI techniques, such as deep learning, requires a significant amount of computational resources, which can hinder their effectiveness on Blockchain. Some experts argue that at the current stage of infrastructure development, it would be better to store data off-chain to ensure that real-time recommendation engines function properly and to minimize latency and costs. The operational risks associated with DLT are still unresolved and will need to be addressed as both the technology and the applications it enables continue to mature.

Smart contracts in DLT-based systems:

- These are decentralized applications that are built and deployed on the blockchain. These applications are made up of self-executing contracts that are written as code on the blockchain ledger. They are designed to automatically execute when certain predetermined events occur, which are also written into the code. This technology has been recognized and discussed by the Organization for Economic Cooperation and Development in 2019.
- They are computer programs that run on the Ethereum blockchain. These programs are designed to determine the functioning and timing of certain actions. Like traditional contracts, they establish a set of rules that are then automatically enforced through the use of code, only when predefined conditions are met.
- They operate autonomously on the network and run on a predetermined schedule, eliminating the need for user control. Users can participate in smart contracts by initiating transactions that trigger specific functions described in the contract.

- They play a crucial role in enabling the disintermediation that distributed ledger technology (DLT) networks can leverage. They offer an important source of efficiency that these networks promise to deliver. By enabling the automation of various actions such as payments and asset transfers based on predefined conditions recorded in code, smart contracts eliminate the need for human intervention. However, despite their potential benefits, the legal status of smart contracts is still uncertain in many areas. They are not yet widely recognized as legal contracts. This lack of clarity raises concerns about enforceability and financial protection when it comes to smart contracts. Furthermore, auditing the code of these contracts requires additional resources from market participants who want to ensure the legitimacy and trustworthiness of the underlying processes.

Self-learning smart contracts and DLT governance

According to the researchers, AI-powered smart contracts have the potential to create self-regulating chains. In the future, AI could be used to forecast and automate processes within "self-learning" smart contracts, similar to the application of reinforcement learning AI techniques. AI can extract and analyze real-time information from systems and incorporate that data into smart contracts. As a result, smart contract code could automatically adapt, eliminating the need for human intervention in chain governance. This would lead to the establishment of fully autonomous and self-regulating decentralized chains.

Decentralized autonomous organizations (DAOs) are entities that operate autonomously on a blockchain, and while they have already been established, incorporating AI-based techniques could enhance their functionality. For example, AI

could provide real-time data to the code, allowing it to calculate the optimal course of action. Self-learning smart contracts that integrate AI would play a crucial role in expanding the capabilities of blockchain logic. These contracts would learn from the blockchain's past experiences, adapt or introduce new rules, and govern the overall functioning of the blockchain.

Currently, most DeFi projects are run by DAOs that possess certain centralized elements, such as on-chain voting by token holders and off-chain consensus. These elements involving human intervention could be subject to regulation. However, by integrating AI into DAOs, it is possible to enhance decentralization and decrease the relevance of traditional regulatory approaches.

Utilizing AI in building fully autonomous chains presents significant hurdles and uncertainties for both users and the ecosystem at large. In such environments, the execution of decisions and the operation of systems would be entrusted to AI smart contracts instead of human involvement, giving rise to crucial ethical concerns (Butijn, 2023). Moreover, implementing automated mechanisms to instantly deactivate the model is particularly challenging in such decentralized networks, which is a major issue faced by the DeFi space as well.

The inclusion of artificial intelligence (AI) in blockchains has the potential to benefit decentralized finance (DeFi) applications by streamlining processes and improving efficiency in the delivery of various financial services. For example, the integration of AI models can enable personalized recommendations for users on different products and services, facilitate online data-driven credit scoring, and offer data-driven investment and trading advisory services. Likewise, the application of reinforcement learning in blockchain-based processes opens up more possibilities for AI in DeFi. It is worth noting that the incorporation of AI into DeFi can expand the capabilities of

distributed ledger technology (DLT) use cases by introducing new functionalities. While it is important to recognize that the introduction of AI may not completely transform the underlying business models in these applications.

AI for ESG investing

ESG ratings[1]Data can differ significantly between various ESG rating providers due to the use of different frameworks, measures, key indicators and metrics, as well as subjective judgement and weighting of subcategories. This disparity in ratings is further exacerbated by the lack of necessary tools, such as consistent data, comparable metrics and transparent methodologies, which are crucial to inform decision-making in the market. The importance of data becomes even more crucial when analyzing non-financial aspects of company performance related to sustainability issues. However, concerns remain regarding the quality of ESG data, including gaps in data availability, potential inaccuracies and the lack of comparability across different providers.

Artificial intelligence and big data have the potential to revolutionize ESG investing by providing a means to assess company data, non-trading data and the consistency and comparability of ratings. The use of AI can significantly improve decision-making by reducing subjective and cognitive biases that often arise from traditional analysis methods, as well as minimizing noise in ESG data and leveraging unstructured data. Natural language processing (NLP) can specifically analyze large amounts of unstructured data, such as geolocation and social media information, to perform sentiment analysis and identify patterns and relationships. These analyses can

[1] The ESG score is determined by collecting and analyzing data related to different aspects and evaluating them based on a predefined scale. The ESG score evaluates a company's achievements and actions in different areas, including, but not limited to, emissions control, protection of human rights, and promotion of sustainable purchasing practices.

then be used to assign quantitative values to qualitative sustainability parameters, using advanced AI techniques.

Recent years have seen the rise of alternative ESG rating providers, with these providers offering AI-powered ratings to facilitate a more impartial external assessment of companies' sustainability performance. By harnessing the power of AI, these rating providers aim to address the problem of greenwashing, which occurs when companies adopt superficial sustainability measures while continuing with their conventional business strategies. By using AI, these providers can uncover crucial insights into companies' sustainability practices and actions that would otherwise not be easily accessible. This innovative approach has immense potential to improve transparency and accountability in the business world.

There is empirical data available to support the use of alternative AI-based ESG ratings. This evidence suggests that these ratings have several key benefits compared to traditional approaches. These advantages include a higher level of standardization, a more democratic aggregation process, and the use of rigorous real-time analytics. However, these methods are unlikely to completely replace traditional models in the future. Instead, they have the potential to work alongside traditional ESG rating approaches, providing additional insights to investors about undisclosed information about the entities being rated.

It is therefore important to recognize that reputable companies themselves can use AI to potentially obscure their sustainability performance. By leveraging AI techniques, these entities can gain a deeper understanding of their operations and accurately identify areas that need to be strategically highlighted in terms of disclosure to improve their environmental, social and governance (ESG) ratings. This allows them to manipulate

their ESG ratings by emphasizing these specific areas, creating a distorted picture of their overall sustainability performance.

Emerging risks from AI/ML/Big Data use: risk mitigation tools

The expansion and diversification of AI/ML technology in financial markets has generated numerous challenges and risks that require careful attention from market participants, industry professionals, and policymakers. These challenges can be observed at various levels, including the data level, the model and business level, as well as the societal and systemic level.

In the rapidly expanding field of AI in finance, there are several challenges that deserve careful attention and consideration. These challenges include data management and concentration, the potential for bias and discrimination, the need for explainable AI models, ensuring the robustness and resilience of AI systems, establishing effective governance and accountability frameworks, addressing regulatory concerns, and managing occupational risks and skills. To mitigate these risks, it is important to explore potential strategies and solutions.

Data and its management

Data plays a central role in all ANN applications driven by AI, machine learning models and big data presents numerous possibilities to improve efficiency, reduce costs and increase customer satisfaction by offering superior services and products (Moujahid, 2016).

The risks that arise when using big data in AI-powered ANN in the financial sector stem from several factors, including the quality of the data used, concerns about data privacy and confidentiality, cybersecurity threats, and fairness and equity considerations.

A key risk is the potential for unintentional bias and discrimination against certain groups of people, which can occur when data is misused or when inappropriate data is used in models, such as in credit underwriting. In addition to consumer protection concerns in the financial space, the use of big data and machine learning models can also raise competition concerns, particularly if there is a high concentration of market providers.

<u>The representativeness and relevance of the data</u>

One of the key aspects of big data, commonly referred to as one of the four "Vs," is veracity, referring to the level of uncertainty surrounding the reliability and accuracy of big data. This uncertainty can arise from a number of factors, such as questionable sources, inadequate data quality, or the inappropriateness of the data being used. In the big data realm, the veracity of observations can be influenced by specific behaviors observed on social media platforms, the presence of noisy or biased data collection systems, such as sensors or the Internet of Things (IoT). As a result, the veracity of big data may not be sufficient to prevent or mitigate disparate dynamic impacts, further complicating its reliability and usability.

The concept of data representativeness and relevance is more significant when it comes to AI applications (ANN) than data veracity. Data representativeness focuses on whether the data used provides a comprehensive and balanced representation of the population being studied, including all relevant subgroups. This is particularly important in financial markets to ensure that certain groups of traders are not over- or under-represented, leading to more accurate model training.

In the context of credit scoring, data representativeness can help promote financial inclusion among minority groups. On the other hand, data relevance refers to the degree

to which the data used accurately describes the phenomenon under study without including misleading information. For example, when evaluating credit scores, it is essential to carefully assess the relevance of information about the behavior or reputation of individuals or legal entities before incorporating it into the model. However, evaluating the data set on a case-by-case basis to improve accuracy and appropriateness can be cumbersome due to the huge volume of data involved, which could undermine the efficiency gained from implementing AI.

Privacy and confidentiality of data

The use of data in AI systems raises several concerns regarding data protection and privacy due to its volume, prevalence, and continuous flow. Aside from typical concerns about the collection and utilization of personal data, the field of AI introduces additional complexities. For example, AI's ability to draw inferences from extensive data sets can create compatibility issues. Traditional privacy practices such as "notice and consent" may not be practical for ensuring privacy protection in machine learning models.

There are also challenges related to data connectivity and the transfer of data across borders. This emphasizes the importance of data connectivity in the financial industry and the critical need to have the ability to aggregate, store, process and transmit data internationally while implementing appropriate safeguards and governance standards.

The process of merging multiple data sets can present numerous advantages for those who are new to working with big data (Harrington, 2018). By combining data from diverse sources, people can get a broader and more complete picture of the information they are analyzing. This is particularly beneficial when dealing with databases that have

been collected under different conditions, such as different populations, regulations, or sampling methods. By bringing together these diverse data sources, new analytical opportunities arise that would not have been possible by examining each data set individually. However, it is important to note that merging heterogeneous data sets also introduces certain challenges and complexities. For example, combining data from different settings can lead to confounding factors, biases in sample selection, and biases that arise when comparing different populations.

The presence of cybersecurity risks, hacking risks, and other operational risks in the field of digital financial products and services has a direct impact on data privacy and confidentiality (Brynjolfsson et al., 2019). While the use of AI technology does not introduce new avenues for cyber breaches, it has the potential to amplify existing vulnerabilities. This amplification can occur through various means, such as the connection between falsified data and cyberattacks, leading to the emergence of new attacks that can alter the functionality of the algorithm by introducing manipulated data into its models. Furthermore, AI can also facilitate the modification of already existing attacks.

The sharing and use of consumers' financial and non-financial information is becoming more frequent, often without their full understanding or consent. While obtaining informed consent is the legal basis for using data, consumers may not always be aware of how their information is handled or where it is used, leading to potential gaps in their understanding and consent. The increased use of advanced tracking methods to monitor online activities and data sharing by third-party vendors further increases these risks. Thus, also data sets that are collected without direct customer input, such as geolocation or credit card transaction data, are particularly susceptible to potential breaches of privacy policies and data protection laws.

The industry is suggesting new methods to protect consumer privacy when calculating confidentiality. One approach is to create and use custom synthetic datasets for machine learning purposes. Another approach is to use privacy-enhancing technologies (PETs), which aim to maintain the general characteristics of the original data while keeping individual data samples confidential. PETs encompass techniques such as differential privacy, federated analysis, homomorphic encryption, and secure multi-party computation. Differential privacy, in particular, offers stronger privacy guarantees and enables more accurate calculations compared to synthetic datasets. The advantage of these techniques is that models trained on synthetic data perform as well as those trained on real data. Traditional data anonymization methods, on the other hand, do not offer strong privacy guarantees, especially considering the inferences made by AI models.

The use of big data in AI-based models has the potential to expand the scope of what is considered sensitive information. These models have the ability to accurately identify individual users by efficiently analyzing a variety of factors, such as facial recognition technology and inferred data, such as customer profiles. By combining this information with other data sources, AI models can make inferences about user characteristics, such as gender, or even re-identify individuals from anonymized databases by cross-referencing them with publicly available information. This process leads to the attribution of sensitive information to specific individuals. Furthermore, the increased dimensionality of ML datasets, which allows for an unlimited number of variables to be considered compared to traditional statistical techniques, increases the likelihood of including sensitive information in the analysis.

Regulators have therefore rekindled the focus on privacy and data protection due to the increasing digitalization of the economy. One such example is the European Union's General Data Protection Regulation (GDPR), which seeks to enhance consumer

protection and rebalance the power dynamics between businesses and individuals. The overall aim is to empower consumers and foster transparency and trust in how businesses handle consumer data. Safeguarding consumer data and privacy is a core principle outlined in the G20-OECD High Level Principles on Financial Consumer Protection. Furthermore, the Monetary Authority of Singapore is committed to upholding fairness, ethics, accountability and transparency in the use of artificial intelligence in the financial sector, with a particular emphasis on protecting individuals' personal data.

The financial sector faces challenges in improving data governance for businesses due to the perception of fragmentation in regulatory and supervisory responsibility regarding data. There is uncertainty about which institutions should implement the most effective data governance practices, including areas such as data quality, definitions, standardization, architecture and de-duplication. This issue becomes even more complex when cross-border activities are considered.

The field of data use in economics is undergoing a transformation due to the widespread adoption of machine learning models in the financial industry (Levenstein & Suslow, 2006). As a result, a few companies specializing in alternative data sets have emerged to meet the growing demand for data to serve as the basis for AI techniques. However, these data set brokers operate with minimal oversight and transparency, raising concerns about the legality of financial service providers' acquisition and use of their data. Furthermore, the rising compliance costs associated with regulations designed to protect consumers may have a profound impact on the economics of big data use in the financial market. This, in turn, will influence how financial market providers approach the use of AI and big data.

Chapter 3

Data and competition in AI-based financial services

Advances in AI have the potential to create competitive advantages that could negatively impact efficient and competitive markets, as consumers may have limited ability to make informed decisions if there is a high concentration of market providers. The use of AI may give certain financial services providers an advantage over smaller competitors who may not have the resources to adopt these technologies. In addition, uneven access to data and the dominance of a few large BigTech companies in obtaining big data could make it difficult for smaller players to compete in the market for AI-based products and services.

The risks of concentration and dependence on a few dominant players are heightened by the potential for network effects, which could result in the emergence of new players that have a significant impact on the entire system. BigTechs, in particular, exemplify this risk, and the fact that they operate outside regulatory boundaries further complicates the challenges associated with this issue. This is primarily due to the way in which big tech companies access and use data, which is further amplified by the use of artificial intelligence techniques to generate profits from that data. In addition, we are witnessing a growing influence of a small number of alternative data providers in the database industry, which could lead to a concentration of power in that market.

When it comes to entering the AI market, smaller companies may encounter data-related hurdles, as they may need to invest in expensive tools such as advanced data mining software and machine learning technology, as well as physical infrastructure such as data centers. Investments are more cost-effective for larger companies due to economies of scale. Additionally, algorithms need to access a wide range of data from

diverse sources to identify new relationships and patterns. Since smaller companies without the necessary resources or a presence in multiple markets may struggle to develop algorithms that can effectively compete with established players, they may face barriers to entry that hinder their ability to succeed in the AI market.

The presence of healthy competition in the market for AI-based financial products and services is crucial for providers to fully leverage the advantages of this technology, particularly in trading and investment (Descamps et al., 2021). The use of third-party or outsourced vendor models can help determine the advantages of these tools for the companies that implement them, but it also has the potential to create one-sided markets and encourage herding behavior among financial consumers. Thus, financial professionals may begin to adopt similar trading and investment strategies, resulting in convergence within the industry.

Tacit collusion: risks

The implementation of AI-based models on a large scale could raise concerns about competition as it allows for the possibility of tacit collusion without any formal agreement or human interaction (Caforio, 2023). Tacit collusion refers to a situation where competitors independently decide on strategies to maximize their own profits, leading to a non-competitive outcome. In simpler terms, the use of algorithms makes it easier for market participants to maintain profits above the competitive level without explicitly colluding, replacing explicit collusion with tacit coordination.

Although tacit collusion typically occurs in markets that are transparent and have a limited number of participants, there is evidence to suggest that collusion becomes more manageable and observable in digital markets involving algorithms. These digital markets are characterized by a high level of transparency and frequent interaction.

The ability of self-learning and deep learning AI models to adapt and learn dynamically has the potential to increase the risk that these models recognize and adjust to the behavior and actions of other market participants or AI models. This could lead to collusive outcomes being reached without any human intervention, and even without the AI model itself being aware of it. While such collusion does not necessarily violate competition laws, it raises concerns about how to address and regulate the model and its users through enforcement measures.

Bias and discrimination

AI methods have the ability to mitigate discrimination by humans in various interactions or amplify prejudice, unfair treatment, and discrimination in the financial services arena. By entrusting the algorithm with the responsibility of decision-making, people using AI-based models can avoid the inherent biases associated with human judgment. However, it is essential to recognize that the adoption of AI applications also presents the possibility of introducing bias or discrimination due to the potential reinforcement of existing biases found in the data. This can occur through training models using biased data or identifying misleading correlations.

Using faulty or inappropriate data in neural networks can lead to incorrect or biased decision-making (Tan, 1997). When poor-quality data is used, biased or discriminatory decisions can be made in two ways. Machine learning models that are trained on inadequate data can produce inaccurate results, even when good-quality data is input. Similarly, machine learning models trained on high-quality data can still produce questionable results if fed inadequate data, despite being well-trained algorithms. Consequently, well-intentioned ML models can unintentionally produce biased conclusions that discriminate against certain protected groups. Using incorrect,

inaccurate (such as mislabeled or incomplete), or fraudulent data in machine learning models poses the risk of "garbage in, garbage out," where the quality of the model output is highly dependent on the quality of the data.

Biases may also exist in the data used as variables, and because the model is trained on data from external sources that may have already incorporated certain biases, it continues to retain these historical biases. Likewise, biased or discriminatory decisions made by machine learning models are not necessarily intentional and can occur even with high-quality, well-labeled data. This can happen through inferences and proxies, or because it is difficult to identify correlations between sensitive and non-sensitive variables within large databases. Since big data encompasses vast amounts of information that reflect society, AI-based ANNs have the potential to perpetuate existing biases present in society and reflected in these databases.

The involvement of humans in AI-based decision-making processes is crucial to detect and rectify any bias that may be present in the data or in the model design. Furthermore, humans play a vital role in interpreting and explaining model outputs, although the feasibility of achieving this to its fullest extent is still uncertain. The human factor is essential both at the data entry and system query stages, and it is important to approach model outputs with a certain level of skepticism to minimize the potential risks of biased results or decision-making.

ML model design and auditing play a crucial role in ensuring model robustness and minimizing potential bias. If AI/ML models are not intelligently designed and controlled, they can unintentionally amplify existing biases and make it even harder to detect discrimination. Conducting model and algorithm audits that compare model outputs to benchmark data sets can help prevent unfair treatment and discrimination.

It is essential that scoring systems can be tested by users and supervisors to ensure fairness and accuracy. Tests can also be run to see if protected classes can be inferred from other attributes in the data, and various techniques can be used to identify and rectify bias in ML models. It is also important to govern AI/ML models and assign responsibility to humans involved in the project to protect potential borrowers from unfair bias. When assessing bias, it is critical to avoid comparing machine learning-based decision making to an ideal unbiased state and instead use realistic benchmarks comparing these methods to traditional statistical models and human-based decision making, as both approaches have their limitations and potential biases.

Explainability

One of the main challenges facing ML models is the difficulty in breaking down the model's output and understanding the factors that contribute to its decision-making process. This concept, known as "explainability," refers to the ability to justify or rationalize the decisions and outcomes produced by the model. AI-based models are inherently complex due to the nature of the technology used, and the intentional concealment of the inner workings of these models by market players further increases the lack of explainability. Simply having access to the underlying code is not enough to fully understand the mechanics of the model, especially considering the widespread lack of technical knowledge among end consumers. This problem is further exacerbated by the mismatch between the complexity of AI models and the cognitive capabilities of humans in terms of reasoning and interpretation.

The lack of trust that users and supervisors have in AI applications is due to the inability to understand and explain how machine learning models work. In the field of finance, AI-driven approaches have become increasingly complex and opaque, making it

difficult for people to understand how these models make decisions. Even if the underlying mathematical principles of these models can be explained, they still lack a clear and explicit explanation of their knowledge.

This lack of transparency therefore undermines trust between financial consumers and regulators, especially in critical financial services. To address this issue, improving the explainability of AI applications is critical as it can help maintain trust in the industry. From an internal control and governance perspective, it is important to ensure that AI models have a minimum level of explainability. This allows a modelling committee to thoroughly analyze the model and feel confident in its implementation.

Thus, the absence of explainability may not be consistent with existing regulations that require understanding and disclosure of the underlying logic. For example, regulations may require comprehensive understanding and explanation of algorithms throughout their lifespan. Other policies may give individuals the right to receive an explanation of decisions made by algorithms, along with information about the logic involved, such as the GDPR in the EU19, which applies to credit decisions and insurance pricing.

Another example is the potential use of ML to calculate regulatory requirements, such as risk-weighted assets (RWA) for credit risk. Current standards require that these models be explainable or at least subject to human oversight and judgment, as described in the Basel Framework for calculating RWAs for credit risk.

The use of ML-based models in financial markets could pose a significant risk if regulators do not closely monitor their lack of explainability. This lack of transparency makes it difficult for both financial firms and supervisors to anticipate how these models will impact markets. This is particularly worrying because AI technology has the

potential to introduce or amplify systemic risks, such as the increased likelihood of herding behavior and strategy convergence among users of generic models provided by third-party vendors.

Without a deep understanding of how these models work, users have limited ability to predict their impact on market conditions or identify whether they are contributing to disruptions. Furthermore, users are unable to adapt their strategies in the face of poor performance or market stress, which can lead to increased market volatility and periods of illiquidity. This can further exacerbate events such as flash crashes. Lack of a clear understanding of model mechanics also creates risks of market manipulation, such as spoofing and tacit collusion between market participants.

Financial professionals in the field of financial markets who use AI-based ANNs are facing increased scrutiny regarding their ability to explain how their models work. This increased attention has led many market participants to focus on improving the explainability of these models. In doing so, they hope to better understand how the models behave both under normal market conditions and times of stress, as well as effectively manage the associated risks.

While achieving explainability by design, i.e., building explainability into the AI mechanism itself, is a challenging task. This is primarily due to several reasons: first, the general public may have difficulty understanding the underlying logic of the model; second, some models, such as certain neural networks, are inherently complex and cannot be fully understood; and third, fully disclosing the mechanism would involve disclosing the intellectual property behind the model (Motta, 2023).

The issue of explainability in relation to AI has sparked a thought-provoking debate about how the level of explainability required for AI differs from that needed for

other complex mathematical models in the financial sector. One concern is that AI applications may be seen as more demanding, leading to a higher burden of explainability compared to other technologies. This potential disparity could have a detrimental effect on innovation within the field. Rather than focusing solely on the mathematical potential of AI models, it is crucial that committees prioritize the analysis of the inherent risks that these models may expose businesses to, determining whether these risks can be effectively managed.

Financial service providers need to strike the right balance between model explainability and accuracy/performance in order to navigate the trade-off between the two. It is crucial that these providers have a certain level of understanding of how the model operates and the underlying logic it follows to avoid being perceived as “black boxes”. By having this knowledge, financial service providers can comply with regulatory obligations and build trust with consumers. In certain areas such as Germany, models that lack some degree of explainability are not accepted.

It is important to note that there is no single principle or approach that can fully explain ML models, and the level of explainability will vary depending on the specific context. When assessing the interpretability of a model, it is necessary to consider the question being asked and the predictions made by the model. Furthermore, it is critical to understand that ensuring the explainability of the model does not automatically guarantee its reliability. To effectively align explainability with the public, a shift in focus towards “risk explainability” is necessary. This means placing more emphasis on understanding the potential risks associated with using the model, rather than focusing solely on the methodology behind the model. The UK Information Commissioner’s Office has recently provided guidance on using five contextual factors (scope, impact, data used, urgency and audience) to assess the type of explanation required.

The auditability of algorithms

The use of "black box" models in financial services, such as lending, presents challenges when it comes to regulatory transparency and auditing. These models are complex and difficult to decompose, making it impossible to understand the underlying drivers of their output. This lack of explainability hampers the ability to perform audits and limits the supervisor's understanding of the model's decision-making process. In some areas, laws and regulations require auditability and transparency, which can be difficult for ANNs to achieve. The ability to follow an audit trail depends on the interpretability of the model, which is often limited in AI models. As the decisions made by these models are already non-linear and their interpretability is limited, it is crucial to find ways to improve the explainability of AI outputs while maintaining accountability and strong governance in AI-based systems.

There are ongoing research efforts in both academia and industry aimed at improving the understandability of artificial intelligence (AI) applications and making machine learning (ML) models more accessible for examination before and after deployment.

The disclosure

The OECD AI-powered ANN Principles emphasize the importance of transparency and responsible disclosure when it comes to AI systems. This means that it is crucial for people to have a clear understanding of AI-based results and to have the ability to question or challenge them. To address the issue of opacity of algorithm-based systems, it is proposed to implement transparency requirements. This would involve providing clear information about the capabilities and limitations of the AI system. The purpose of separate disclosure is to inform consumers about the use of AI in the delivery

of a product and their interaction with an AI system rather than a human, as in the case of robo-advisors. By having this information, customers can make informed decisions and choose between different competing products.

So far, there is no widely accepted standard on the amount of information that should be disclosed to investors and financial consumers, as well as on the proportionality of such information. Market regulators suggest that the level of transparency should vary depending on the type of investor (retail or institutional) and the area of application (front or back office). These regulators argue that suitability requirements, such as those applied to the sale of investment products, could help firms assess whether potential customers have a comprehensive understanding of how the use of AI affects the provision of the product or service.

Requirements for financial firms to document operational details and design features of financial models existed even before the emergence of AI. Some regulators are now using documentation of algorithm logic as a means to ensure that model outputs can be explained, traced and repeated.

The European Union, for example, is contemplating implementing requirements to disclose documentation on methodologies, processes, programming, and training techniques used in the development, testing, and validation of AI systems, including documentation on the algorithm itself (OECD, 2021). The US Association for Computing Machinery (USACM) Public Policy Council has proposed a set of principles prioritizing transparency and auditability in the use of algorithms, suggesting that models, data, algorithms, and decisions should be recorded to allow for auditing in case of suspected harm. The Federal Reserve's guidance on model risk management also emphasizes the need for detailed documentation of model development and validation, allowing people

unfamiliar with the model to understand its operation, limitations, and assumptions (OECD, 2021).

Financial service providers are facing increasing challenges when it comes to documenting the process of AI-based models used for monitoring purposes. This difficulty arises from the complex nature of these models, which makes it difficult to explain how they work and subsequently document them. This challenge is not limited to the size of the service provider, as even smaller providers face the same obstacles.

In response, some areas have proposed a two-pronged approach to overseeing AI models:

- The first aspect is analytical and combines source code and data analysis to document AI algorithms, predictive models and data sets, preferably following standardized methods.
- The second aspect is empirical and uses techniques that provide explanations for individual decisions or the general behavior of the algorithm.

This is achieved through the use of challenger models, which are used to compare against the model being tested, as well as benchmarking data sets selected by auditors.

Furthermore, aside from the difficulties associated with explaining AI-based models, there is the added complexity of setting numerous parameters that affect model performance and outcomes. The parameterization process can be considered "arbitrary" and subjective, as it is often based on intuition rather than thorough validation and is heavily influenced by the individual designing the model. While revealing the chosen parameters might alleviate some of the problems, explaining how these parameters interact with the model still poses a significant challenge.

Robustness and resilience of AI models

It is crucial for AI systems to operate effectively, safely, and reliably at every stage of their existence, and it is imperative to continuously assess and mitigate any potential risks they may pose. To improve the robustness of AI systems, it is essential to diligently train models and thoroughly evaluate their performance in line with intended objectives.

Training AI models

To account for more complex relationships and non-linearity in the data, it may be necessary to train models using larger datasets. This is because higher-order interactions may be harder to reach and require more data to discover. Therefore, it is crucial to have sufficiently large datasets for training in order to capture non-linear relationships and rare events in the data. However, this presents challenges in practice, as tail events are infrequent and the dataset may not be robust enough to produce optimal results. Furthermore, there is a trade-off, as using increasingly larger datasets to train models risks making them less adaptive, potentially compromising their performance and ability to learn effectively.

The financial system is at risk because the industry has failed to train models on data sets that include rare and unexpected events. This weakens the reliability of AI models in times of crisis and limits their usefulness to stable market conditions. One potential problem is overfitting, where a model performs well on the data it was trained on, but poorly on new, unknown data. To address this, model builders split data into training and validation sets and use the training set to build multiple models with different settings. The validation set is then used to test the accuracy of the models and optimize their parameters. By analyzing the errors in the validation set, the best set of model parameters can be determined.

Previously, scientists believed that the measured performance of validation models provided an unbiased estimate of their overall performance. However, recent studies by Westerhuis et al. (2008) and Harrington (2018) have shown that this assumption is not always accurate. These studies emphasize the importance of having an additional blind test dataset that is not used during the model selection and validation process. This test dataset is essential to obtain a more reliable estimate of the model's ability to generalize. The validation processes mentioned in these studies involve more than just retrospectively testing the model on historical data to assess its predictive capabilities. They also aim to ensure that the results obtained from the model are reproducible.

Artificially generated synthetic datasets are being used as test sets for validation purposes. These datasets present an intriguing alternative due to their ability to provide unlimited amounts of simulated data. Furthermore, they offer a potentially more cost-effective approach to improving predictive accuracy and strengthening the resilience of machine learning models. This is especially beneficial in situations where obtaining real data is difficult and expensive. In certain cases, regulatory bodies, such as those in Germany, require the evaluation of AI model results within test scenarios defined by supervisory authorities.

Ongoing monitoring and validation of models throughout their lifecycle is crucial to effectively managing the risks associated with any type of model. Model validation is conducted after model training and serves to confirm that the model has been implemented correctly and is performing as intended. It encompasses a number of processes and activities intended to ensure that models align with their design objectives and business purposes, while ensuring their robustness. This involves identifying potential limitations and assumptions and assessing their potential impact. The

validation process should encompass all aspects of the model, including input data, processing, and reporting. It applies to both internally developed models and those obtained from external or third-party sources. Validation activities should be conducted continuously to monitor known model limitations and identify new ones, particularly during periods of economic or financial stress that may not be reflected in the training data set.

Continuous testing of ML models is extremely important to detect and rectify any drift in the models called "model drifts". These drifts can occur in the form of concept drifts or data drifts. Concept drifts refer to situations where the statistical characteristics of the target variable being analyzed by the model undergo changes, thereby altering the fundamental concept that the model aims to predict. For example, as time passes, the understanding of fraud may evolve due to the emergence of new methods used for illegal activities. This evolution in the definition of fraud would lead to concept drift.

Data drift occurs when the statistical characteristics of the input data undergo alterations, which impacts the model's ability to make accurate predictions. A noteworthy example of such data drift is the pronounced shift in consumer sentiments and inclinations towards e-commerce and digital banking. These changes, which were not accounted for in the original training dataset, can lead to a decrease in model performance.

Continuous monitoring and validation of machine learning (ML) models plays a crucial role in preventing and addressing drift. By implementing standardized procedures for this monitoring process, we can improve model resilience and determine whether any tuning, redevelopment, or replacement is necessary. It is of utmost importance to establish an efficient architecture that facilitates rapid retraining of models

with updated data, especially when data distribution changes, as this helps mitigate potential risks associated with model drift.

In addition to continuously monitoring and evaluating the code or model being used, certain regulatory bodies have implemented a requirement to include "kill switches" or other automated control mechanisms that trigger alerts in high-risk situations. Kill switches serve as an example of such control mechanisms, as they can quickly disable an AI-based ANN if it deviates from its intended purpose. As an example, in Canada, companies are mandated to incorporate "override" features that can automatically shut down the system or allow for remote shutdown if deemed necessary. These kill switches must undergo rigorous testing and ongoing monitoring to ensure that companies can rely on them if the need arises.

Existing risk management functions and processes specifically designed for AI-based models to address emerging risks and unintended consequences need to be enhanced. To ensure the effectiveness of models, performance testing under extreme market conditions is critical. This is essential to prevent the emergence of systemic risks and vulnerabilities that may arise during stressful times. However, it should be noted that the data used to train these models may not fully capture the effects of market stress conditions or changes in various factors such as exposures, activities or behaviors.

Consequently, this limitation could negatively impact model performance. Furthermore, since these models are new, their ability to effectively address risks under changing financial conditions has not yet been tested. To mitigate this, it is important to incorporate a multitude of scenarios for testing and back testing purposes. By considering different market behaviors and trends, there is hoped to minimize the possibility of underestimating risks in such scenarios.

Research has indicated that the explainability of a system in a way that humans can easily understand can have a substantial impact on how users perceive its accuracy, regardless of the true observed accuracy. According to the OECD (2018), when explanations are provided in a way that is less understandable to humans, users are less inclined to accurately assess the accuracy of a technique that is not based on easily understood principles.

Meaningless learning

The convergence of causal inference and machine learning has emerged as a burgeoning field of study, as indicated by the rapid growth of research in this area. While pattern recognition systems lack the ability to understand cause-effect relationships, understanding such relationships is a fundamental aspect of human intelligence. Consequently, there is a growing recognition among deep learning researchers of the importance of such research and they are incorporating it into their studies. However, it is important to note that this particular area of research is still in its nascent stages.

Users who employ machine learning models may run the risk of misinterpreting meaningless correlations observed in activity patterns as causal relationships. This can lead to questionable model results. It is crucial to go beyond mere correlation and dig deeper into causality to understand the circumstances under which a model might fail. This understanding will allow us to determine whether the observed pattern will remain predictive over time.

Likewise, causal inference plays a vital role in replicating a model's empirical results in new settings, environments, or populations, thereby ensuring the external validity of the model's results. The ability to transfer causal effects learned from a test dataset to a new dataset, where only observational studies can be conducted, is referred

to as transportability. This concept is fundamental to the utility and robustness of machine learning models. Supervisors may find it beneficial to gain insight into causal assumptions made by users of AI models to better assess potential associated risks.

It is crucial to thoroughly evaluate the results of AI models, and human judgment plays a vital role in this process, particularly when it comes to determining causality. Without a healthy dose of skepticism or caution, relying solely on correlation identified by AI-based models can lead to biased or inaccurate decision-making, as causality may not necessarily be present. Research has shown that models are prone to acquiring suboptimal strategies if they do not consider human advice, even in cases where human decisions may be less accurate than the models' own skills.

<u>The example of the COVID-19 crisis</u>

While AI-based ANNs are designed to adapt and learn from new data over time, they can struggle to handle unique, unforeseen events such as the COVID-19 crisis. These events are not accounted for in the data used to train the models, making it difficult for them to operate effectively. AI-driven trading systems, which rely on dynamic models trained from historical data, are often successful as long as the market environment remains consistent with the past.

While a survey of UK banks indicates that approximately 35% of them experienced a negative impact on the performance of their machine learning models during the pandemic (OECD, 2018). This can be attributed to significant changes in macroeconomic variables caused by the pandemic, such as rising unemployment and changes in mortgage lending, which required recalibrating both machine learning and traditional models. Unforeseen events such as the pandemic disrupt the continuity of data sets, leading to model drifts that undermine the predictive capabilities of these models.

Tail events refer to unexpected occurrences that lead to unforeseen changes in the behavior of the target variable, thereby affecting the accuracy of model predictions. These events also cause previously unrecognized alterations in the underlying data structure and patterns of the dataset used for model training, all due to changes in market dynamics during such events. Since these changes are not accounted for in the original dataset, they are likely to result in a decrease in model performance. To address this, future synthetic datasets created for model training could incorporate similar tail events, along with data from the COVID-19 period, in order to retrain and distribute updated models.

Therefore, it is crucial to engage in continuous model testing using validation data sets that encompass extreme scenarios. Furthermore, it is vital to continuously monitor any drift in the models. This is essential to minimize potential risks that may arise during periods of stress or uncertainty. It is worth mentioning that reinforcement learning-based models, where the model is trained using simulated conditions, are predicted to exhibit superior performance during rare and unforeseen events that pose extreme risks. This is because such models are comparatively easier to train by incorporating conditional scenarios, even those involving extraordinary and unprecedented market trends that have not been observed in the past.

Chapter 4

Governance of AI systems

Establishing robust governance structures and transparent accountability mechanisms is crucial when deploying ANN in critical decision-making scenarios, such as determining access to credit or allocating investment portfolios. It is imperative that organizations and individuals involved in the development, implementation or operation of AI systems take responsibility for ensuring their effective and accountable functioning. As stated by the OECD, strict measures are needed to enforce accountability. Likewise, the European Commission (2020) emphasizes the importance of human oversight throughout the lifecycle of AI products and systems to protect against potential risks and biases.

Currently, financial market players using AI rely on existing governance and oversight mechanisms when using these technologies. This is because AI-based algorithms are not considered fundamentally different from traditional algorithms. Current governance frameworks that apply to models can serve as a basis for developing or adapting to AI activity, as many of the considerations and risks associated with AI are also relevant to other types of models.

By implementing explicit governance frameworks that clearly establish lines of responsibility for the development and oversight of AI-based systems throughout their entire lifecycle, from development to deployment, existing AI-related operational arrangements can be further strengthened. These internal governance frameworks may include minimum standards or guidelines on best practices and approaches to implementing these guidelines. The establishment of internal model committees plays a key role in establishing model governance standards and the processes that financial

service providers follow when creating, documenting and validating models of any type, including AI-based machine learning models.

Current model governance frameworks have not yet considered the unique challenges posed by AI models, which have a transient existence and undergo frequent changes. The problem lies in adapting existing model governance processes to accommodate more advanced AI models that have the ability to rebuild themselves in short periods of time. One solution to address this problem is to preserve the data and code used in the model, allowing the generation of replicas of the model's inputs and outputs based on past dates. However, it is important to note that many ML models are non-deterministic, meaning that even with the same input data, there is no guarantee that the exact same model will be produced.

Incorporating desired outcomes for consumers into a governance framework is of paramount importance, and this should be accompanied by an assessment of whether and how these outcomes are achieved through the use of AI technologies. When it comes to advanced deep learning models, there may be concerns about who controls the model, as AI could unintentionally act in a way that goes against the best interests of consumers. For example, biased outcomes in credit underwriting, as mentioned above, could be a potential consequence. Furthermore, the autonomous behavior exhibited by certain AI systems throughout their lifecycle may lead to significant changes to the product, which could affect its safety. Consequently, a new risk assessment may be necessary in such cases, as highlighted by the European Commission in 2020.

Ultimate responsibility for AI-based systems lies with the executive and management levels of the financial services provider. They must establish a comprehensive approach to managing model risk and ensuring it is within acceptable levels. In addition, other roles such as engineers, programmers and data analysts, who

have not traditionally been central to supervisory review, may now face increased scrutiny due to their increasing importance in the implementation of AI-based financial products and services.

Therefore, responsibility for AI-related systems may need to extend beyond senior management and the board to professionals responsible for programming, model development and use of the system. It is crucial that these technical functions have a mechanism to provide services to customers and effectively explain these models to senior management and the board. In some areas, a third-party audit may be required to validate the performance of the model in accordance with its intended purpose. Strong governance also involves thorough documentation of model development and validation.

Typically, financial services providers employ similar procedures for developing, documenting, and validating machine learning (ML) models as they do for conventional statistical models.

The implementation of best practices in model governance has been in place since the adoption of conventional statistical models for credit and consumer finance determinations. It is imperative for financial institutions to ensure that models are built using appropriate data sets and refrain from incorporating certain data into the models. It is also critical to avoid the use of surrogate data that can potentially discriminate against protected groups. Rigorous testing and validation of models, sometimes conducted by independent validators, is also essential. Furthermore, when models are used in live operations, it is vital to ensure that input data aligns with data used during the model development phase. Adequate audit trails and documentation are maintained to track various aspects such as implementation, design, and production decisions.

Model governance frameworks also emphasize the importance of monitoring models to ensure that they do not produce results that indicate unequal treatment. Therefore, it is critical to have the ability to understand the reasoning behind the model output. In the financial services sector, organizations establish model governance committees or model review boards to develop, authorize, and oversee the implementation of model governance procedures.

Model validation is a crucial aspect of various procedures that involve the use of retained data sets. In addition to this, there are other conventional procedures such as examining the consistency and reliability of inputs, outputs, and parameters. As AI adoption becomes more prevalent in the financial industry, the establishment of internal committees to oversee these processes is expected to become increasingly common. Furthermore, these committees are likely to undergo enhancements in their roles and powers to accommodate the intricate nature of AI-based models. It is important to note that the frequency and methodologies employed for model validation in the context of AI-based models should be different from those applied to linear models.

Artificial intelligence is also being used for regulatory technology (RegTech) purposes. To ensure effective model governance, financial services firms are actively working to improve automated procedures that monitor and regulate the data used by operating models. In addition, they are also focusing on improving automated systems that monitor and evaluate the outputs generated by these models.

<u>Outsourcing: Third-party providers</u>

One aspect of the risks involves competitive dynamics, specifically concentration risks. When companies rely on a single third party for their AI needs, there is a risk of becoming overly dependent on that provider. This can create a situation where the

company has limited options and negotiating power, which could lead to higher costs or inferior services. Thus, if the chosen third party experiences financial difficulties or goes out of business, it can disrupt the company's AI operations and cause significant setbacks.

In addition, outsourcing AI techniques can create systemic vulnerabilities, particularly related to increased risk of convergence. Convergence risk refers to the potential for multiple systems or processes to become interconnected and dependent on one another. By outsourcing AI techniques to third parties, companies are introducing an external element into their operations, which can increase the complexity and interconnectedness of their systems. This can make the company more vulnerable to failures or disruptions in the third-party AI infrastructure, which could lead to operational disruptions or compromise data security.

There are additional risks that need to be considered when outsourcing AI techniques to third parties. These risks can be categorized into two main areas: competitive dynamics and systemic vulnerabilities. Outsourcing AI techniques to third parties introduces additional risks beyond the initial benefits. These risks include concentration risks, where companies become overly dependent on a single vendor, and systemic vulnerabilities that arise from increased risk of convergence. It is essential that companies carefully assess and mitigate these risks to ensure the successful implementation and operation of outsourced AI techniques.

Potential concentration risks associated with specific third-party providers may increase when it comes to data collection and management, such as dataset providers, or in the realm of technology provision, such as third-party model providers, and infrastructure, such as cloud providers. As artificial intelligence (AI) models and techniques become more readily available through cloud adoption, there is an increased risk of reliance on outsourced solution providers, creating new challenges in terms of

competitive dynamics and the potential formation of oligopolistic market structures within these services.

Therefore, the use of third-party models has the potential to create convergence risks both at the level of individual firms and at a broader systemic level. This risk is particularly heightened when there is a lack of diversity among third-party models in the market. In times of financial stress, such as those of low liquidity, this convergence risk can lead to herding and instances of illiquidity, which can be detrimental to overall market stability. Equally, the diminishing storage capacity of traditional market makers further exacerbates this problem, as they are unable to provide sufficient liquidity in times of market stress through active market making. Smaller entities are particularly vulnerable to the impact of herding, as they often rely on third parties to handle the development and management of machine learning models due to a lack of internal expertise in this area (OECD, 2021).

Outsourcing AI techniques or the technologies and infrastructure that enable them presents challenges in terms of liability and concentration risks. To effectively manage these risks, it is critical to establish appropriate governance arrangements and contractual modalities, similar to those used in other service sectors. Finance providers must possess the necessary skills to audit and conduct due diligence on services offered by third-party entities. However, an excessive reliance on outsourcing can increase the likelihood of service disruptions, which could have significant systemic impacts on markets. It is therefore imperative to have contingency and security plans in place to ensure that the business can operate smoothly even if any vulnerabilities arise.

Regulatory Considerations

While a significant number of countries have established comprehensive AI strategies, it is worth noting that only a few areas have implemented specific regulations and requirements relating specifically to algorithms and AI-based ANNs. In most cases, oversight and control of machine learning applications is governed by general guidelines for systems and controls. These guidelines typically emphasize the thorough examination and evaluation of algorithms prior to their introduction into the market, as well as the ongoing assessment of their effectiveness and functionality throughout their operational life.

Many areas take a technology-neutral approach when it comes to regulating financial market products, including oversight of risk management, governance and the use of algorithms. However, this approach may face challenges as the innovative use of technology in finance becomes more complex. With advances in artificial intelligence, particularly in areas such as deep learning, existing regulatory frameworks in the financial sector may not adequately address the systemic risks that could arise from the widespread adoption of these techniques.

It should also be noted that certain advanced AI techniques may not conform to current legal or regulatory requirements. This problem arises due to the lack of transparency and explainability of some machine learning models, as well as the ever-evolving nature of deep learning models that are continually being adapted. These factors can potentially create a conflict with existing regulations.

Inconsistencies may also arise in the area of data collection and management. For example, the European Union's General Data Protection Regulation (GDPR) imposes restrictions on storing individual data for a limited period of time. While AI-related regulations might require companies to keep a complete record of the data sets used to train their algorithms for audit purposes, this creates a dilemma as the data sets used to

train these algorithms are often extremely large, leading to practical challenges and costs associated with recording data for monitoring purposes (Klein, 2020).

Certain areas, such as the European Union (EU), have recognized the need to modify or clarify existing laws in specific areas, such as liability, to ensure effective implementation and enforcement of these regulations. The reason behind this need is the lack of transparency in AI systems, which creates challenges in identifying and proving potential violations of laws. This includes legal provisions safeguarding fundamental rights, establishing liability, and allowing for compensation. In the near future, regulators and supervisors may find it necessary to modify regulations and adjust their supervisory approaches to adapt to new realities brought about by the deployment of AI, such as increased concentration and outsourcing.

The regulatory landscape surrounding AI is at risk of fragmenting at several levels, including national, international and sectoral. Industry participants emphasize the need for greater consistency in regulations to ensure that AI techniques can be effectively used across borders. In addition to existing regulations for AI models and systems, numerous principles, guidelines and best practices have been published in recent years. While these resources are seen as valuable in addressing potential risks, there are divergent opinions on their practical utility and the challenges of translating them into effective guidance with real-life examples.

The availability and simplicity of standardized AI tools have the potential to incentivize unregulated entities to offer investment advice or other services without obtaining the necessary certification or license, thereby operating in a non-compliant manner. This phenomenon of regulatory arbitrage is not only observed among large technology companies, but also within their operations, where they use data sets accessible through their core business activities.

Occupational hazards

Financial service providers and supervisors must be technically capable of operating, inspecting AI-based systems and intervening when necessary. Lack of appropriate skills is a potential source of vulnerabilities for both the sector and regulators and supervisors and can lead to potential employment issues in the financial sector. The deployment of AI and big data in finance requires different skills that are possessed by a small segment of financial professionals. In line with the significant investments that will need to be made to develop AI-based models and tools, firms will also need to develop human capital with the skills required to derive value from these technologies and exploit the value of large amounts of unstructured data sources.

From an industry perspective, the deployment of ANNs involves the use of professionals who combine scientific knowledge in the area of AI, computer skills (programming, coding) and experience in the financial sector. While current participants in financial markets have isolated the functions of IT or finance specialists, the widespread use of AI by financial institutions will increasingly depend on, and generate greater demand for, experts who successfully combine financial knowledge with computer expertise (Metaxa et al., 2021). It is important that compliance professionals and risk managers have a proper understanding of how AI techniques and models work in order to be able to audit, monitor, challenge and approve their use. Likewise, senior managers, who are in most cases responsible for the use of these techniques, must be able to understand and follow their development and application.

The widespread adoption of AI and ML by the financial sector may pose some employment challenges. On the one hand, the demand for employees with applicable knowledge in AI methods, advanced mathematics, software engineering and data science

is expected to be significant. On the other hand, executives at financial services firms anticipate that the application of these technologies may lead to potentially significant job losses across the sector. In practice, financial market professionals and risk management experts are expected to gain experience and knowledge in AI in the medium term, as AI models will co-exist with traditional models and until such time as AI becomes mainstream.

Over-reliance on fully automated AI-based systems may lead to a higher risk of service disruption with potential systemic repercussions on markets. If markets relying on such systems face technical or other disruptions, financial services providers should ensure that they are prepared, from a human resources perspective, to replace automated AI systems with well-trained humans who act as a human safety net and are able to ensure that market disruptions do not occur. These considerations are likely to become increasingly important as AI deployment becomes more widespread in markets.

The issue of skills and expertise is becoming increasingly important from a regulatory and supervisory perspective as well. Financial regulators and supervisors may need to keep pace with technology and improve the skills needed to effectively supervise AI-based financial applications. Enforcement authorities may need to be technically capable of inspecting AI-based systems and empowered to intervene when necessary. Training policymakers will also enable them to expand their own use of AI in RegTech and SupTech, an important area of innovation in the official sector.

The use of ANN in finance should be seen as a technology that augments human capabilities rather than replacing them. It could be argued that a "man-machine" combination, where AI informs human judgment rather than replacing it (as a decision aid rather than a decision maker), could allow the benefits of the technology to be harnessed, while maintaining safeguards of accountability and control over the ultimate

decision making. In the current state of maturity of AI solutions, and to ensure that vulnerabilities and risks arising from the use of AI-based techniques are minimized, some level of human oversight of AI techniques remains necessary. Identifying points of convergence where humans and AI are integrated will be critical to the practical application of this combined "man-machine" approach.

Political implications

Political activity around RNA in finance

With the power to revolutionize various industries and the emergence of new risks associated with the implementation of neural networks and their empowering effect on artificial intelligence (AI), this has become an increasingly important focus in policy debates. In May 2019, the Organization for Economic Co-operation and Development (OECD) launched its AI Principles, which mark the first set of globally accepted guidelines for the responsible and ethical use of AI. These principles were formulated by a diverse group of experts from various sectors, ensuring a comprehensive approach to the responsible implementation of trustworthy AI. The breadth of topics covered by the OECD AI Principles and their direct connection to fostering sustainable and inclusive growth make them particularly relevant when considering their application in the realm of global finance.

The Recommendation on AI was officially adopted by the OECD Council during a ministerial-level meeting held on 22-23 May 2019. This important milestone signifies the OECD's commitment to address the challenges and opportunities associated with artificial intelligence (AI) technologies. The OECD AI Principles, which form the core of

this Recommendation, emphasize the crucial role of governments in shaping a human-centered approach to trustworthy AI.

By promoting the use of innovative and trustworthy AI systems, these principles aim to ensure the protection of human rights and the preservation of democratic values. This comprehensive framework serves as a guide for policymakers and stakeholders and offers a roadmap for the responsible development and deployment of AI technologies worldwide.

The Recommendation presents a set of five interrelated principles rooted in ethical values that should guide the responsible management of trustworthy AI. These principles emphasize the importance of AI's contribution to promoting inclusive growth, sustainable development, and the general well-being of both people and the environment.

- Artificial intelligence systems must be built with due regard to the principles of the rule of law, human rights, democratic values and diversity. It is essential that these systems incorporate appropriate safeguards, such as provisions for human intervention where deemed necessary, in order to promote a just society and ensure justice and equality for all.
- To ensure understanding and accountability of AI systems, transparency and responsible disclosure practices are imperative. This allows people to understand the results generated by AI and gives them the opportunity to question or challenge these results.
- AI systems must operate reliably and safely at all times during their existence, and any potential hazards must be constantly assessed and monitored.

- Organizations and individuals that develop, implement or use artificial intelligence systems must be responsible for their correct operation in accordance with the above principles.

The OECD also offers five recommendations to governments:

- To foster the advancement of trustworthy AI, it is important to promote and support public and private investments in research and development, which in turn will foster innovation in this field.
- Promote the development of open and inclusive AI ecosystems supported by advanced digital infrastructures, technologies and efficient mechanisms for data and knowledge sharing.
- Create an enabling policy environment that promotes the implementation of trustworthy and reliable AI systems.
- One way to make a significant impact is to equip people with the necessary AI skills and support workers as they transition to a more equitable future.
- To promote responsible management of trustworthy AI, it is essential that different countries and industries work together and collaborate. By transcending borders and sectors, we can collectively strive to achieve ethical and trustworthy AI practices.

In 2020, the European Commission published a White Paper presenting several strategies and regulations to establish an "AI ecosystem for excellence and trust". This proposal not only outlines specific measures to support the development and adoption of AI in the EU economy and public administration, but also offers potential options for a future regulatory framework for AI.

The White Paper also examines important considerations such as security and liability in the field of AI. The European Commission is also taking practical steps to implement these ideas, including initiatives such as the pilot projects of the EC-funded Infinitech consortium. These projects aim to reduce barriers to AI-driven innovation, improve regulatory compliance and encourage investment in the sector.

The Infinitech project is an ambitious undertaking led by a collaborative consortium of 48 participants from 16 EU member countries. This innovative initiative has received substantial funding from the European Commission's prestigious Horizon 2020 Research and Innovation Programmed. The main focus of the Infinitech project revolves around conducting a wide range of experiments and tests spanning over 20 pilot projects and financial institutions. These tests specifically delve into the field of digital finance, harnessing the transformative power of innovative technologies such as artificial intelligence, big data and the Internet of Things (IoT).

Infinitech offers a wide range of innovative AI-powered products and services. These include a variety of applications such as Know Your Customer (KYC), customer analytics, personalized portfolio management, credit risk assessment, fraud and financial crime prevention, insurance services, and RegTech tools. These tools are specifically designed to incorporate data governance capabilities and ensure compliance with regulations such as PSD2, 4AMLD, and MIFiD II. By leveraging advanced AI technology, Infinitech is able to deliver innovative solutions that enhance customer experiences, improve risk assessment processes, prevent fraudulent activities, and streamline regulatory compliance for businesses in the financial and insurance sectors (Westerhuis et al., 2008).

- Infinitech has carried out numerous pilot projects that serve as shining examples of its innovative approach and commitment to pushing the boundaries in the field.

- An advanced and automated platform has been developed to assess the credit risk of small and medium-sized enterprises (SMEs). This platform uses big data, artificial intelligence (AI) and Blockchain technology to provide accurate credit risk ratings for SMEs.
- Real-time risk assessment in the field of investment banking involves the implementation of a real-time risk monitoring and assessment system that focuses on two commonly used risk metrics, namely VaR (Value at Risk) and ES (Expected Shortfall). This procedure enables a comprehensive assessment of potential risks, providing valuable insights into the potential losses an institution may face. By continuously monitoring and analyzing these risk metrics, investment banks can proactively identify and mitigate potential risks, thereby safeguarding their financial stability and optimizing their investment strategies.
- Customer-centric collaborative data analytics is becoming increasingly important in the financial services industry. An emerging trend in this area is the use of artificial intelligence (AI)-based support tools to improve new customer services. These tools rely on a sophisticated system that facilitates data sharing, incorporates a credit scoring system, and employs anti-money laundering (AML) measures based on semantic technologies. In addition, this system uses distributed ledger technology (DLT) to enable secure and efficient data exchange. By leveraging these advanced technologies, financial service providers can improve their ability to analyze customer data, offer personalized services, and ensure regulatory compliance.
- AI-powered portfolio construction for wealth management, tailored to individual needs, regardless of portfolio size.

- The primary goal of the anti-money laundering monitoring platform is to improve the efficiency of current monitoring practices, such as analytical reporting, risk assessment, and detection tools, by utilizing Big Data processing techniques. By leveraging the power of Big Data, the platform aims to optimize the overall effectiveness of anti-money laundering efforts.
- Real-time cybersecurity analysis is performed on a large amount of financial transaction data, focusing specifically on mobile banking transactions. This analysis incorporates machine learning models and employs advanced analytics techniques to effectively handle the massive influx of data. By doing so, it enables early identification and response to any abnormal activity with appropriate countermeasures.

In 2019, the IOSCO (International Organization of Securities and Exchange Commission) Board of Directors placed particular emphasis on the topic of artificial intelligence (AI) and its potential connection to money laundering. This recognition of the importance of AI continued into the following year, as in 2020, IOSCO published a consultation report specifically addressing the use of AI by market intermediaries and asset managers. The intention behind this report was to present six distinct measures that could help IOSCO members establish appropriate regulatory structures to effectively supervise these intermediaries operating within the market, as well as asset managers employing these advanced technologies.

These aspects include:

- Establishing appropriate governance structures, controls and oversight frameworks to govern the development, testing, use and monitoring of artificial intelligence and machine learning (ML) systems.

- The consultation emphasizes the importance of equipping staff with appropriate knowledge, skills and experience to effectively implement, monitor and challenge AI and ML outcomes.
- To improve the overall robustness and consistency of AI and ML systems, IOSCO emphasizes the need for companies to adopt clear and well-defined processes for development and testing, which allow them to identify and address potential issues before full deployment of AI and ML.
- Finally, the consultation underlines the importance of transparency and disclosure, highlighting the need for companies to provide sufficient information to investors, regulators and other relevant stakeholders about the use of AI and ML technologies in their operations.

Efforts to address the implications of AI in the financial sector have extended to the national level. For example, the French ACPR established a collaborative working group in 2018 bringing together professionals from various financial entities, including business associations, banks, insurers and FinTechs, along with public authorities. The main objective of this group is to facilitate discussions on current and potential applications of AI in the sector, identifying both the opportunities and risks associated with its implementation.

This initiative also aims to address the challenges faced by supervisors in overseeing the adoption of AI in the financial industry. Similarly, in 2019, the Bank of England and the Financial Conduct Authority jointly launched the Public Private Forum on AI, which serves as a platform to engage stakeholders and foster dialogue on the implications of AI in the financial domain (see Box 4.4 for more details).

Similarly, the Russian Federation has made significant strides in developing and regulating AI. In 2019, they enacted a National Strategy specifically dedicated to advancing AI, followed by the introduction of a Concept for regulating AI technologies and robotics in 2020. Furthermore, in 2021, the Russian government passed the Federal Law on Experimental Digital Innovation Regimes, granting the Bank of Russia the authority to approve regulatory sandboxes serving projects involving AI solutions in finance. This legislative move was complemented by the launch of a five-year regulatory sandbox in Moscow in July 2020, under a special Federal Law, specifically designed to facilitate the implementation of AI in the financial sector.

In recent times, various regulatory and policymaking agencies, such as the Comptroller of the Currency, the Federal Reserve System, the Federal Deposit Insurance Corporation, the Consumer Financial Protection Bureau, and the National Insurance Administration Credit Unions, have taken significant steps to address the issue of artificial intelligence (AI) use by financial institutions. This can be seen in their joint initiative, which began on March 31, 2021, in which they requested information and comments on the use of AI, including machine learning, in the financial sector.

The aim of this consultation is to comprehensively assess the potential benefits and risks associated with the implementation of AI in finance. Some of the key concerns highlighted in the consultation include the need for explainability in AI systems, ensuring appropriate use of data and dynamic updating, and addressing potential issues related to intensive lending practices. In addition, the consultation seeks views on how to address the risk of overfitting, mitigate cybersecurity risks, consider fair lending practices, implement effective third-party oversight, and explore other relevant considerations.

On 21 April 2021, the European Commission published a proposal for a regulation that aims to address the potential risks associated with artificial intelligence (AI) and

establish consistent rules for its use across all sectors. As part of this proposal, the creation of the European AI Council is suggested. While the proposal is broad in scope, it imposes the most stringent requirements on high-risk AI applications, such as creditworthiness assessment.

These requirements include the use of comprehensive risk and quality management systems, subjecting the AI system to a conformity assessment, and using high-quality data that is accurate, representative, and complete. Thus, the proposal emphasizes the need for transparency in the use and operation of AI-based applications, the requirement for human oversight by appropriately trained individuals, and the implementation of safeguards such as kill switches or explicit human confirmation of decision-making. It also emphasizes the importance of ensuring the accuracy, robustness, and security of AI systems, conducting post-market monitoring, reporting significant incidents to regulators, and registering the system in a public registry.

Political considerations

The increasing use of artificial intelligence (AI) in the financial services field has the potential to offer substantial benefits to both financial consumers and market participants (França et al., 2021). Not only can it improve the overall quality of services provided, but it can also create efficiencies for financial services providers. While it is critical to recognize that the integration of AI-based applications in the financial industry can also introduce new challenges, such as a lack of transparency and explainability in decision-making processes. There is also the potential for existing risks in financial markets, such as those associated with data management and use, to be further magnified by the adoption of AI technology.

It is crucial that policymakers and regulators prioritize aligning the implementation of AI in the financial sector with the objectives of enhancing financial stability, safeguarding the interests of financial consumers and fostering market integrity and competition. To achieve this, it is imperative to actively identify and mitigate any potential risks that may arise from the use of AI techniques, whilst encouraging and supporting the responsible use of AI. This may involve reviewing and refining existing regulatory and supervisory frameworks to address any perceived inconsistencies or challenges posed by the integration of AI technologies in the financial industry.

The application of regulatory and supervisory measures to AI techniques can be approached in a way that considers the specific context and scale of the application, as well as the potential consequences for people using AI. By adopting a proportionate framework, the aim is to promote the uptake of AI technology while avoiding any undue obstacles to innovation.

It is critical that policymakers pay particular attention to improving data governance within financial sector firms to enhance consumer protection across all aspects of AI implementation in finance. This note highlights several important risks associated with data management, including concerns about data privacy, confidentiality, data concentration, and the potential impact on market competition dynamics.

There is also a risk of unintentional bias and discrimination as a result of data characteristics and trends. The importance of data cannot be questioned, especially in relation to training, testing and validating machine learning models. Furthermore, data plays a critical role in determining the ability of these models to maintain their predictive accuracy during extreme and unforeseen events.

One approach that policymakers could take is to implement specific guidelines or standards for data management in AI-based techniques. These guidelines could cover several aspects such as data quality, ensuring that the dataset used aligns with the intended purpose of the AI model, and implementing safeguards to ensure that the model is robust and free from bias.

To mitigate discrimination risks, it would be beneficial to employ best practices, such as comparing model outputs to established data sets and testing to determine whether protected characteristics can be inferred from other attributes of the data. Another way to minimize bias is to validate the appropriateness of the variables used in the model. It might also be beneficial to develop and use tools to monitor and correct any conceptual bias. In addition, policymakers may want to consider imposing additional transparency requirements on the use of personal data and providing individuals with the option to opt out of the use of their personal data.

Policymakers should consider implementing regulations that require financial service providers to disclose their use of AI techniques and how this may impact customers. It is crucial that financial consumers are fully informed about the use of AI in the products they purchase, as well as the possibility of interacting with an AI system instead of a human representative. This transparency enables consumers to make informed decisions when choosing between different products.

The information disclosed should also provide clear details about the capabilities and limitations of the AI system. To further enhance consumer protection, authorities could also introduce suitability requirements for AI-based financial services, similar to the regulations currently in force for the sale of investment products. These requirements would ensure that financial service providers can accurately assess whether potential

customers have a sufficient understanding of how the use of AI affects the delivery of the product.

The limited transparency and explainability of many advanced AI-based AI models is a key policy issue that remains to be resolved. Lack of explainability is inconsistent with existing laws and regulations, but also with financial service providers' internal governance, risk management and control frameworks. It limits users' ability to understand how their models impact markets or contributes to market disruptions. It can amplify systemic risks related to procyclicality, convergence and increased market volatility through simultaneous buying and selling of large amounts, particularly when using third-party standardized models. More importantly, users' inability to adjust their strategies in times of stress can exacerbate market volatility and lead to episodes of illiquidity during periods of acute stress, aggravating flash crash-type events.

Regulators should consider how to overcome the perceived incompatibility of the lack of explainability in AI with existing laws and regulations. Currently applicable frameworks for model governance and risk management by financial services firms may need to be updated and/or adjusted to address the challenges posed by the use of AI-based models. Supervisors' focus may need to shift from documenting the development process and the process by which the model arrives at its prediction to the behavior and outcomes of the model, and supervisors may wish to look at more technical ways to manage risk, such as adversarial model stress testing or outcome-based metrics.

Despite recent advances in improving AI explainability from low levels, explainability remains at the core of perceived lack of trust from users and supervisors around AI applications. While current discussions tend to focus on improving explainability as the sole mechanism to promote trust, other checks and balances may

need to be introduced to ensure that decision making based on AI models works as intended.

Policymakers may consider requiring clear governance frameworks for models and attribution of responsibility to humans to help build trust in AI-based systems. Financial services providers may need to establish explicit governance frameworks that designate clear lines of responsibility for the development and oversight of AI-based systems throughout their lifecycle, from development to deployment, in order to reinforce existing arrangements for AI-related operations.

Governance frameworks for internal models may need to be adjusted to better capture the risks arising from the use of AI, as well as to incorporate intended outcomes for consumers along with an assessment of whether and how those outcomes are achieved using AI technologies (Westerhuis et al., 2008). Adequate documentation and audit trails of the above processes can assist supervisors in monitoring this activity.

The provision of greater assurances by financial firms on the robustness and resilience of AI models is critical as policymakers seek to guard against the build-up of systemic risks and will help AI applications in finance gain confidence. Testing the performance of models under extreme market conditions may be necessary to prevent systemic risks and vulnerabilities that may arise in times of stress.

Introducing automatic control mechanisms (such as kill switches) that trigger alerts or disable models in times of stress could help mitigate risks, although they expose the firm to new operational risks. Backup plans, models and processes should be in place to ensure business continuity in the event that models fail or act unexpectedly. Regulators could also consider additional or minimum buffers if banks were to determine risk weights or capital based on AI algorithms.

Frameworks for proper training, retraining and rigorous testing of AI models may need to be introduced and/or strengthened to ensure that ML model-based decision making is working as intended and in compliance with applicable rules and regulations. Datasets used for training should be broad enough to capture non-linear relationships and tail events in the data, even if synthetic, to improve the reliability of such models in unforeseen times of crisis. Continuous testing of AI models is indispensable to identify and correct model drift.

Regulators should strongly advocate for continuous monitoring and validation of AI models, as these activities play a crucial role in risk mitigation. By emphasizing the importance of these practices, regulators can help improve the resilience of models and effectively address any deviation from their intended performance. Developing standardized procedures for monitoring and validation would be particularly beneficial, as it would establish best practices that can be universally adopted. Such procedures would also allow for the identification of models that require adjustment, refurbishment, or replacement. To ensure transparency and accountability, it is essential to separate model validation from its development process and to thoroughly document all relevant information. Furthermore, the frequency of testing and validation should be determined based on the complexity of the model and the importance of the decisions it influences.

The importance of human involvement in decision-making becomes particularly relevant in situations where high-value decisions, such as credit decisions, have a significant impact on consumers (Caforio, 2023). To foster trust in these systems, regulatory authorities could consider implementing processes that allow customers to challenge the results of AI models and seek solutions. The General Data Protection Regulation (GDPR) serves as an example of such a policy, as it gives individuals the right to request human intervention and express their concerns if they wish to question

decisions made by algorithms (EU, 2016). Furthermore, clear and transparent communication by government entities about their expectations can further enhance trust in the use of AI applications in the financial sector.

Policymakers need to consider the increasing complexity of AI technology and consider whether they will need to allocate resources to keep up with developments. Investing in research can help address issues related to understanding and unintended consequences of AI techniques. In addition, it is important to invest in skills for both financial sector participants and policymakers so that they can stay informed about technological developments and engage in interdisciplinary discussions at various operational, regulatory and supervisory levels.

A solution to balance model predictability and explainability, as well as meeting legal and regulatory transparency requirements, could be to foster closer collaboration between IT professionals and traditional finance experts. This could involve bridging the gap between disciplines such as deep learning and symbolic approaches, which involve human-created rules, to improve the explainability of AI-based approaches. It may also be necessary for law enforcement authorities to possess technical capabilities to inspect AI-based systems and have the authority to intervene when necessary, while benefiting from the use of AI by implementing RegTech/SupTech applications.

The role of policymakers plays a crucial role not only in supporting innovation in the sector, but also in ensuring adequate protection of consumers and financial investors, as well as maintaining fair, orderly and transparent markets for these products and services. Policymakers may need to adjust and enhance their existing measures to effectively address the risks associated with the use of AI. An important aspect of this is to clearly communicate the adoption of AI and the safeguards put in place to protect the system and its users, which can help build trust and promote the implementation of these

innovative techniques. Given the easy cross-border provision of financial services, it is essential to foster and maintain a multidisciplinary dialogue between policymakers and the sector, both at national and international levels.

Conclusions

Over the past few decades, financial markets have undergone significant changes thanks to the emergence of advanced communication and trading platforms, which have allowed a greater number of investors to access the markets, leading to a transformation of traditional capital market theory and an improvement in financial analysis methods.

Researchers have long been intrigued by the prediction of stock returns, which typically involves examining the relationship between publicly available fundamental information from the past and future returns of stocks or indexes. This approach challenges the efficient market hypothesis, which holds that all relevant information is quickly incorporated into stock prices, making it impossible to predict future returns. While there is conflicting evidence suggesting that markets may not always be fully efficient, leaving room for the possibility of predicting future returns with better-than-probability outcomes.

Considering the research conducted, it is evident that there is evidence supporting the predictability of stock market returns using publicly available information such as time series data on financial and economic variables. The studies highlight the importance of variables such as interest rates, monetary growth rates, changes in industrial production and inflation rates in predicting a portion of stock returns.

However, it is important to note that most of these studies rely on simple linear regression assumptions, despite the lack of evidence supporting a linear relationship between stock returns and financial and economic variables. Since there is a considerable amount of residual variance in actual stock returns compared to predictions made by regression equations, it is possible that the use of nonlinear models can account for this

residual variance and provide more accurate forecasts of stock price movements (Nagesha et al., 2016).

Due to the prevalence of linear assumptions in current modeling techniques, it becomes essential to consider a financial analysis method that incorporates nonlinear analysis of embedded financial markets. Although nonlinear regression can be performed, most of these techniques require the specification of a nonlinear model before determining parameter estimates. However, neural networks present a nonlinear modeling technique that can overcome these challenges (Odom & Sharda, 1990).

Neural networks offer a unique approach that requires no pre-specification during the modeling process, as they autonomously learn the inherent relationship between variables. This is particularly valuable in securities investing and other financial areas where assumptions abound and little is known about the underlying processes that determine asset prices. In addition, neural networks provide the advantage of flexible architectural choices, learning algorithms, and validation procedures.

Literature

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The journal of finance, 23(4), 589-609.

Assad, S., Clark, R., Ershov, D., & Xu, L. (2020). Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market. https://www.econstor.eu/bitstream/10419/223593/1/cesifo1_wp8521.pdf

Botta, M., & Wiedemann, K. (2020). To discriminate or not to discriminate? Personalized pricing in online markets as exploitative abuse of dominance. European Journal of Law and Economics, 50, 381-404.

Botta, M., & Wiedemann, K. (2019). Exploitative behaviors in digital markets: time for a discussion after the Facebook decision. Journal of European Competition Law & Practice, 10(8), 465-478.

Brynjolfsson, E., Collis, A., & Eggers, F. (2019). Using massive online choice experiments to measure changes in well-being. Proceedings of the National Academy of Sciences, 116(15), 7250-7255.

Brown, Z.Y., & MacKay, A. (2021). Competition in pricing algorithms (No. w28860). National Bureau of Economic Research.

Butijn, B.J. (2023). Introduction to Advanced Information Technology. Advanced Digital Auditing, 15.

Caforio, V. (2023). Algorithmic Tacit Collusion: A Regulatory Approach (15 Competition Law Review 9), Available at SSRN: https://ssrn.com/abstract=4164905 or http://dx.doi.org/10.2139/ssrn.4164905

Cheng, T.K., & Nowag, J. (2023). Algorithmic Predation and Exclusion. U. Pa. J. Bus. L., 25, 41.

Descamps, A., Klein, T., & Shier, G. (2021). Algorithms and competition: the latest theory and evidence. Competition Law Journal, 20(1), 32-39.

França, RP, Monteiro, ACB, Arthur, R., & Iano, Y. (2021). An overview of deep learning in big data, image, and signal processing in the modern digital age. Trends in Deep Learning Methodologies, 63-87.

Harrington, P. (2018), Multiple Versus Single Set Validation of Multivariate Models to Avoid Mistakes, Taylor and Francis Ltd.. http://dx.doi.org/10.1080/10408347.2017.1361314.

Klein, T. (2020), "(Mis)understanding Algorithmic Collusion", Competition Policy International. https://www.competitionpolicyinternational.com/misunderstanding-algorithmic-collusion/

Levenstein, M.C., & Suslow, V.Y. (2006). What determines cartel success? Journal of economic literature, 44(1), 43-95.

Metaxa, D., Park, JS, Robertson, RE, Karahalios, K., Wilson, C., Hancock, J., & Sandvig, C. (2021). Auditing algorithms: Understanding algorithmic systems from the outside in. Foundations and Trends® in Human–Computer Interaction, 14(4), 272-344.

Motta, M. (2023). Self-preferencing and foreclosure in digital markets: theories of harm for abuse cases. International Journal of Industrial Organization, 102974.

Moujahid, A. (2016). A practical introduction to deep learning with caffe and python. Access http://adilmoujahid.com/posts/2016/06/introduction-deep-learning-python-caffe

Nagesha, K., Chandar, K.R., & Sastry, V. (2016). Prediction of dust dispersion by drilling operation using artificial neural networks. Int. J. Prev. Control Ind. Pollut, 1, 1-13.

Odom, M.D., & Sharda, R. (1990, June). A neural network model for bankruptcy prediction. In 1990 IJCNN International Joint Conference on neural networks (pp. 163-168). IEEE.

OECD (2023). Algorithmic Competition, OECD Competition Policy Roundtable Background Note. www.oecd.org/daf/competition/algorithmic-competition-2023.pdf

OECD (2021), Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers. https://www.oecd.org/finance/artificial-intelligence-machine-learningbig-data-in-finance.htm.

OECD (2018). Personalized pricing in the digital era.

Saurwein, F., Just, N., & Latzer, M. (2015). Governance of algorithms: options and limitations. info, 17(6), 35-49.

Schrepel, T. (2020). The fundamental unimportance of algorithmic collusion for antitrust law. SSRN.

Tan, C.N. (1997). An artificial neural networks primer with financial applications examples in financial distress predictions and foreign exchange hybrid trading system. Bond University, 50-78.

Westerhuis, JA, Hoefsloot, HC, Smit, S., Vis, DJ, Smilde, AK, van Velzen, EJ, ... & van Dorsten, FA (2008). Assessment of PLSDA cross validation. Metabolomics, 4, 81-89.

This edition of "Data science and artificial intelligence: Finance, policy and governance" was completed in the city of Colonia del Sacramento in September 2024.

EST. 2021 EMC
EDITORIAL MAR CARIBE

DATA SCIENCE AND ARTIFICIAL INTELLIGENCE:

Finance, policy and governance

www.editorialmarcaribe.es
Colonia del Sacramento
Uruguay
2024

ISBN: 978-9915-9706-5-3